Essays on the Foster–Hart Measure of Riskiness and Ambiguity in Real Options Games

Inauguraldissertation

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.) durch die Fakultät für Wirtschaftswissenschaften der Universität Bielefeld

vorgelegt von

Diplom–Mathematiker

Tobias Hellmann

Institut für Mathematische Wirtschaftsforschung

Universität Bielefeld

Oktober 2015

Erstgutachter:	Prof. Dr. Frank Riedel
Zweitgutachter:	Jun.-Prof. Dr. Jan–Henrik Steg
Drittgutachter:	Prof. Dr. Herbert Dawid

Acknowledgment

First and foremost, I would like to thank my advisor Frank Riedel for his continuous support of my doctoral studies. I have learned a lot through the fruitful input he gave as my advisor and co–author. His motivation, support and guidance were essential for the whole research process and the completion of the present thesis.

In addition, I would like to express my gratitude to my second advisor Jan–Henrik Steg. He always took time for discussion and I am especially grateful that he shared some inside with me into the game theoretic foundations of stochastic timing games.

I am also very grateful to Jacco Thijssen for being a great co–author and for his hospitality during my research stay in York.

I thank my fellow members of the IMW for the enjoyable social activities as well as for a lot of fruitful discussions. Also the administrative assistance provided by the chairman of our graduate school Herbert Dawid, the managing director Ulrike Haake and our secretaries Bettina Buiwitt–Robson, Sabine Heilmann, Karin Borchert and Diana Grieswald–Schulz has been a great support for me during my studies.

Personally, I would like to thank my parents who supported me throughout my entire life with all their love and care.

Finally and most importantly, I would like to thank my wife Mona. Her support and love were undeniable the bedrock on which this thesis was built. Throughout my whole time as a doctoral student, Mona has been a great companion who always encouraged me and helped me to make progress with the thesis.

Summary

In this thesis, we apply two different concepts of uncertainty to economic problems. Following Knight (1921), we distinguish between risk, where the uncertainty can be captured by a single probability measure, and Knightian uncertainty (or Ambiguity), where we use a set of priors because it is not possible to assign a single probability measure to the underlying uncertainty.

In the first part of the thesis, we extended the Foster–Hart measure of riskiness to both general gambles and dynamic frameworks. Foster and Hart (2009) proposed an operational measure of riskiness for discrete random variables. Their defining equation has no solution for many common continuous distributions. We show how to extend consistently the definition of riskiness to continuous random variables. For many continuous gambles, the risk measure is equal to the worst–case risk measure, i.e. the maximal possible loss incurred by that gamble. For many discrete gambles with a large number of values, the Foster–Hart riskiness is close to the maximal loss. We give a simple characterization of gambles whose riskiness is or is close to the maximal loss.

We also extend the Foster–Hart risk measure to dynamic environments for general distributions and probability spaces, and we show that the extended measure avoids bankruptcy in infinitely repeated gambles.

In the second part of this thesis, we study a two–player investment game with a first mover advantage in continuous time with stochastic payoffs. One of the players is assumed to be ambiguous with max–min preferences over a strongly rectangular set of priors. We develop a strategy and equilibrium concept allowing for ambiguity and show that equilibira can be preemptive (a player invests at a point where investment is Pareto dominated by waiting) or sequential (one player invests as if she were the exogenously appointed leader). Following the standard literature, the worst–case prior for the ambiguous player if she is the second mover is obtained by setting the lowest possible trend in the set of priors. However, if the ambiguous player is the first mover, then the worst–case prior can be given by either the lowest or the highest trend in the set of priors. This novel result shows that "worst–case prior" in a setting with κ–ambiguity does not equate to "lowest trend".

Contents

Chapter 1

General Introduction

Uncertainty is a major ingredient of life. Except for the promises made by dubious visionaries, no human being can predict what the future will bring. Yet, everyday decisions have to be made, the outcomes of which heavily depend on the uncertain future development of certain underlying factors.

In Economics in particular, one is interested in the behavior of agents under uncertainty. Wrong decisions along with poorly developing markets may have a huge impact on the whole society as the recent financial crisis so drastically demonstrated.

The present thesis applies two different aspects of uncertainty to several economic questions. In the first part, we consider so–called risky situations, i.e. situations in which the underlying stochastic is captured by a single probability measure. A tool that measures the underlying risk in monetary form is provided. This is done by extending an existing risk measure developed in Foster and Hart (2009) to both continuous distributions and dynamic frameworks.

In the second part of the thesis, the uncertainty of the market is understood in a much broader sense. We allow for different possible dynamics driving the underlying stochastic. This form of uncertainty is called Knightian Uncertainty or Ambiguity in the literature. We introduce the concept of Knightian Uncertainty to a game of two firms competing for the same irreversible investment project. The distribution of the underlying stochastic is assumed to be ambiguous.

1.1 Risk measures

A major issue that all financial institutions have to be concerned with is the measurement of the risk that their future endeavors entail. One of the risks that banks, for instance, have to take into account is the risk that debtors might not be able to settle the money they borrowed. For this reason a debtor has to pay interest the amount of which depends on how likely it is that she might not be able to pay back the loan. With these interest rates, the creditor receives compensation for the risk and the loss of opportunity of instantaneous consumption.

Further examples of risk that financial institutions have to handle are "market risk" (i.e. the risk resulting from moving market prices) and "operational risk" (i.e. the risk resulting from failed internal processes, mistakes made by people or external events).

In 1974, the presidents of the central banks of the ten biggest economies were worried that banks did not have enough equity to overcome critical times. For this reason, the "Basel Committee on Banking Supervision" was founded, which developed in 1988 a framework or set of rules for financial institutions. This set of rules eventually led to guidelines and laws in more than 100 countries. Among other things, these guidelines stipulate to banks how to measure and ensure the risky components of their portfolio. Revising and improving these guidelines led to a more sophisticated set of rules called Basel 2 in 2004. However, in the course of the recent financial crisis, it became evident that Basel 2 still needed drastic improvements. The result was published as Basel 3 in 2010.

The measure on which the whole risk quantifying process is based in the Basel 2 and Basel 3 reports is called Value at Risk. The Value at Risk for a given confidence level α determines the maximal possible loss of a financial position if one ignores losses that occur with a probability smaller than or equal to α.

Many papers, however, discuss the shortcomings of this measure, see among others Artzner et al. (1999) and Föllmer and Schied (2004). One of the undesirable properties of Value at Risk is that it is not subadditive. This means that the risk of a portfolio can be higher than the accumulated risk of the individual positions in it. For this reason, an investor might want to split her portfolio into several single accounts in order to reduce the risk. Such a behavior contradicts one of the main techniques of reducing risk proposed in the finance literature, namely diversification. Another shortcoming is that Value at Risk could give incentives to take the risk of very high losses as long as they occur with a small enough probability. As Value at Risk does not recognize losses which occur with a sufficiently small probability, an investor might want to take high risks if they come (as it is often the case) with the possibility of great gains. If these small probability events, however, take place, they could lead to bankruptcy.

The hedge fund manager David Einhorn referred to Value at Risk as "an airbag that works all the time, except when you have a car accident".[1] In his popular scientific book, Riedel (2013) even saw in the use of Value at Risk one of the reasons for the emergence of the recent financial crisis. It is therefore crucial for the research community to come up with some more sophisticated measures that are both applicable for the financial industry and able to overcome the negative properties of Value at Risk.

In the field of mathematical finance, risky assets are described by random variables on a probability space which take values in the real numbers. Those numbers denote the possible gains and losses of the positions. A (monetary) risk measure is defined as a real valued function on a space of such assets. That means, a risk measure assigns to every position a real number denoting the risk of it.

This is a very general notion for a risk–quantifying method, of course. Risk measures should certainly satisfy several desirable properties. For instance, if an asset X has higher payments than an asset Y in every state of the world, the risk of X should not be higher than the risk of Y. This monotonicity property, along with positive homogeneity, subadditivity and translation invariance are proposed by Artzner et al. (1999) as desirable properties. Risk measures which satisfy these properties are called coherent risk measures and are formally

[1] See Einhorn (2008).

defined as follows: let A be the set of all assets, that is the set of all real valued functions on a probability space.

Definition 1 *(Artzner et al. (1999)) A risk measure* $\rho : A \to \mathbb{R}$ *is called coherent if it fulfills*

- *Monotonicity: For all* g *and* $h \in A$*, if* $g \leq h$ *then* $\rho(h) \leq \rho(g)$*.*
- *Positive Homogeneity: For all* $\lambda \in \mathbb{R}_+$[2] *and all* $g \in A$*,* $\rho(\lambda g) = \lambda\rho(g)$*.*
- *Subadditivity: For all* g *and* $h \in A$*,* $\rho(g + h) \leq \rho(g) + \rho(h)$*.*
- *Translation Invariance: For all* $g \in A$ *and all* $\alpha \in \mathbb{R}$*,* $\rho(g+\alpha) = \rho(g) - \alpha$[3]*.*

In Artzner et al. (1999), the research is done on a finite probability space. Delbaen (2002) extended the theory of coherent risk measures to arbitrary probability spaces and discussed the connection between coherent risk measures and the theory of cooperative games. Föllmer and Schied (2004) and Frittelli and Rosazza Gianin (2002) weakened the axioms positive homogeneity and subadditivity and replaced them with convexity, defining in that way the more general class of convex risk measures.

1.2 The Foster–Hart Measure of Riskiness for Continuous Distributions

An important task for supervising agencies is to ensure that banks, or more general financial players do not go bankrupt. The complicated network on financial markets between banks, insurance companies and firms demonstrates the dependence of the financial players on each other. The bankruptcy of even just one of these players may affect other players that are connected with her immensely. The bankruptcy of Lehman Brothers in 2008 is a good example

[2] Define $\mathbb{R}_+ := \{x \in \mathbb{R} | x \geq 0\}$.

[3] Artzner et al. (1999) defined the condition of translation invariance as $\rho(g + \alpha \cdot r) = \rho(g) - \alpha$, where r is the "total rate of return on a reference instrument". That means r is some kind of interest rate on a risk-free investment. In our case we say $\alpha = 1$.

for the impact one player may have on the whole financial industry. In the course of Lehman Brothers' bankruptcy, many banks got in trouble and had to be bailed out. As described above, the commonly used Value at Risk is not an appropriate measure for that purpose.

An approach that aims to find a tool that helps prevent bankruptcy was proposed by Foster and Hart (2009). They developed a new risk measure, the Foster–Hart measure of riskiness, which yields a minimal capital reserve an agent has to have in order to avoid bankruptcy in the long run. More precisely, they showed that a decision maker who is offered a gamble g_t in every discrete time period t avoids bankruptcy with probability one if and only if her wealth W_t at time t exceeds the assigned risk $R(g_t)$. A gamble is defined as a real valued discrete random variable, that has a positive expected value and where losses occur with positive probability.

Formally, R is defined as the unique positive solution to

$$E \log \left(1 + \frac{g_t}{R(g_t)} \right) = 0. \tag{1.1}$$

The Foster–Hart measure of riskiness satisfies all properties given in Definition (1) except for translation invariance.[4]

Foster and Hart (2009) considered only discrete distributed gambles. Most of the financial options and portfolios are, however, priced by using continuous distribution. For instance, the famous Black–Scholes option pricing model developed by Black and Scholes (1973) and Merton (1973) relies on lognormal distributed prices.

It seems therefore natural and important to extend the Foster–Hart measure to continuous distributions. This is studied in the second chapter of this thesis. One of our main findings is that the defining equation (1.1) does not admit a solution for arbitrarily continuous distributions. In fact, there exists a whole class of distribution for which the defining equation does not possess a

[4]For the purpose of the Foster–Hart measure, it is indeed reasonable to drop the cash invariance property. See chapter VI.D in Foster and Hart (2009) for a more detailed discussion about the connection between the Foster–Hart measure and the class of coherent measures.

solution. We show, however, that for this class, a reasonable extension is given by the maximal loss of these gambles. This extension is motivated by the fact that for discrete distributions which converge to such continuous distributions, the riskiness numbers converge to the maximal loss. Furthermore, given a sequence of discrete gambles which converge to a continuous distributed gamble X for which a solution to equation (1.1) exists, we show that the respective sequence of riskiness numbers converge to the solution $R(X)$ of equation (1.1). We therefore argue that the robust extension of the Foster–Hart measure of riskiness is given by either *the (unique) solution to equation* (1.1) if it exists or *the maximal loss of the given gamble*, otherwise.

This extension is now applicable to all continuous distributed gambles. The extended risk measure is applied to many commonly used continuous distributions such as the lognormal distribution and uniform distribution in section 2.3.

1.3 Dynamic Foster–Hart Measure of Riskiness

An important justification for our extension is the fact that the no–bankruptcy result of Foster and Hart (2009) carries over to the extended version. To show this, we need to embed the Foster–Hart measure of riskiness into a dynamic framework. The analysis in chapter 2 and in Foster and Hart (2009) is done in a static framework. In the third chapter of this thesis, a dynamic framework for our extended version of the Foster–Hart measure of riskiness is introduced.

Dynamic measurement of risk plays an important role when it comes to incorporate the arrival of new information about the assets over time. The arrival of new information yields the opportunity to quantify the risk of these positions in a more precise way. If we want to measure the risk of an asset that has its payments in, say, a year from now, it would be very likely that in six months new information about the asset is revealed or the market situation has been changed in a way that the assigned risk should be adjusted accordingly.

Since the late 1990s, many works have dealt with the incorporation of new information to risk measures, see among others Wang (1999) and Karatzas and Cvitanic (1999). The information over time is modeled by a filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$. A dynamic risk measure is then defined as a family of conditional risk measures $(\rho_t)_{t\in\mathbb{N}}$, where every $\rho_t(X)$ determines the risk of a financial position given the information $\mathcal{F}_t$ at time t. That means every $\rho_t(X)$ is a $\mathcal{F}_t-$measurable random variable. Wang (1999) shows some basic desirable properties a dynamic risk measure should satisfy. In Riedel (2004) and Detlefsen and Scandolo (2005), the class of coherent risk measures is extended to a dynamic framework.

Another useful property is the notion of time–consistency. Suppose, for instance, you know that tomorrow in every state of the world a position X will never have a higher risk than a position Y. It would seem rather unreasonable if today X is assigned to have a greater risk than Y. A dynamic risk measure that is not time–consistent might yield an inconsistent behavior of a decision maker. In terms of risk, the decision maker would prefer position Y to X today, even if she knows that tomorrow, her preferences about these assets will be the other way around. Time–consistency is exactly the condition which avoids such an unreasonable behavior. Time–consistency, in particular for dynamic coherent risk measures, is studied in Riedel (2004), Cheridito et al. (2004), Detlefsen and Scandolo (2005), and Cheridito and Kupper (2011).

The dynamic version of the Foster–Hart measure of riskiness is proposed in chapter 3 as a family of conditional Foster–Hart risk measures. These conditional measures are defined in a very similar way to the extended Foster–Hart measure of chapter 2, but now conditioned on the information given at the respective time. To verify the existence and uniqueness of this family of random variables, however, is not trivial and to prove it requires some work which is done in detail at the end of the chapter.

Besides the benefits a dynamic measurement of risk delivers, the dynamic Foster–Hart measure also justifies our extension of the Foster–Hart measure. It is shown that the no–bankruptcy result carries over to the extended version of the Foster–Hart measure. That means in particular that for some gambles, a wealth equal to the maximal loss of the gamble suffices to accept the gambles without taking the risk of going bankrupt. This result seems surprising at first sight, as in the original work of Foster and Hart (i.e. for discrete gambles), a wealth that is strictly greater than the maximal loss is required to overcome the possibility of bankruptcy. However, in contrast to discrete distributed gambles, the maximal loss occurs for continuous distributed gambles with a probability of zero (and losses that are close to the maximal loss with very small probability), which makes a big difference for the purpose of staying solvent.

Despite these nice properties, one shortcoming of the dynamic Foster–Hart measure of riskiness is demonstrated in section 3.4. We show that it is not time–consistent. In fact, we give an example where two different gambles will have the same risk tomorrow. Yet, today, one gamble is assigned to have a higher risk than the other. This contradicts the time–consistency property.

1.4 Knightian Uncertainty

We can think of many situations where uncertainty about the future is greater than considered in the first two chapters of this thesis. Often times it is not possible to give an exact stochastic description of future events. Think, for instance, about the weather forecast or betting on a soccer game. Whereas the probabilities of the result of a coin–flip can be exactly determined (and we speak of risk), the determination of the probability that the sun will shine next Sunday for more than 5 hours is fairly vague (and we speak in this case of uncertainty). In fact, most of the weather forecasting portals indeed assign such probabilities, yet the weather is such a complex system that depends on many different factors which make it impossible to assign an exact probability for the given event.

Similar situations can be observed in financial markets. In classical financial markets, the dynamics of the assets are assumed to be known. However, the only thing that is indeed known is the past development of the assets. Often times, banks compute from these past data just one particular stochastic or dynamic underlying these assets. Yet, as the recent financial crisis demonstrates, there are far more events that are not included in the model delivered by past data, effecting the future development of the assets.

Uncertainty that cannot be described by a single probability distribution was first discussed by Frank Knight in his famous book *Risk, Uncertainty and Profit* (Knight (1921)). Knight described there why he distinguished between risk and uncertainty: "*Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated. [...] The essential fact is that risk means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character; and there are far–reaching and crucial differences in the bearings of the phenomena depending on which of the two is really present and operating. [...] It will appear that a measurable uncertainty, or risk proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all.*"

This unmeasurable uncertainty was afterwards named after Frank Knight as "Knightian Uncertainty".

Also in decision theory the approach of subject expected utility and the sure thing principle introduced by Savage (1954) were questioned by the famous Ellsberg paradox brought up by Ellsberg (1961). Roughly speaking, the sure thing principle states that if two acts are equal on a given event, it does not matter, for ranking these events in terms of preferences, to what they are equal on that event. Ellsberg, however, indicates that the sure thing principle does not hold if agents are uncertain about the probability with which payoff–relevant events occur.

Knightian uncertainty, or as it is often called, Ambiguity, was rigorously formalized in a decision theoretical framework by Gilboa and Schmeidler (1989) and Bewley (2002).[5] Gilboa and Schmeidler and Bewley allow for a whole set of probability measures which possibly describe the dynamics of the underlying in order to account for the uncertainty in the market. Gilboa and Schmeidler (1989) compare different options under the probability measure that delivers the worst–case, i.e. the probability measure delivering the minimal expected utility. Under Bewley preferences, a decision maker prefers an option over another if and only if it is unanimously preferred under all probability measures.

[5] Bewley wrote his article in the late 1980s. It was published more than ten years later in 2002.

A framework that has as a special case the Gilboa Schmeidler max–min approach and allows for subjective attitude towards ambiguity is given by the so–called smooth ambiguity model developed by Klibanoff et al. (2005). They allow, essentially, the decision maker to assign a subjective probability to a possible probability measure representing the decision maker's subjective opinion on how likely it is that this probability measure is the "right one".

A further important model which allows one to measure ambiguity aversion is given in Maccheroni et al. (2006). This so–called variational preference model adds to the multiple priors model of Gilboa and Schmeidler a convex function that captures the attitude of a decision maker towards ambiguity. Similar to the smooth ambiguity model, it is possible to distinguish between attitude towards risk and ambiguity.

1.5 Knightian Uncertainty and Real Options Games

In the second part of the thesis, we adopt ambiguity into a real option game.

As a "real option", we understand the right but not the obligation firms might have to invest in a certain irreversible project. The valuations of those projects can be done by using well–known pricing methods from the option pricing literature.

The real option theory goes back to the 1980s, when McDonald and Siegel (1986) studied an irreversible investment option of a firm which might enhance the profit flow. The main problem studied has to do with discussing the optimal point in time of investment. It is shown that in a stochastic framework the net present value rule fails to be the optimal decision rule. In fact, investment should be done later than it is suggested by the NPV rule.

Considering not only a single firm (monopoly), but rather two (or more) firms, evokes the need to include game theoretical components into the model. Many works have dealt with this combination of real option theory and game theory, such as Dixit and Pindyck (1994), Huisman (2001), Mason and Weeds (2010) and Thijssen et al. (2012), just to name a few. Most of the literature concentrates either on preemption games such as Smets (1991), Huisman and Kort (1999) and Boyer et al. (2004), or (much less studied) war of attrition games, see Hoppe (2000) and Murto (2004). In a preemption game, firms have a first–mover advantage and profit from being the first to invest, whereas in a war of attrition game, it is beneficial not to be the first to invest due to a second–mover advantage. A framework that connects preemption and attrition is proposed in Steg and Thijssen (2015). Such models are solved backward in time by first computing the follower value function, assuming the competitor has already invested. In a second step, the leader value function is determined. Using these functions, equilibrium results can be obtained.

All papers mentioned so far, however, assume the firm(s) to be certain about the right distribution of the profit flow. As discussed above, this is quite a restrictive assumption. Indeed, it is very unlikely that the market uncertainty can be captured by a single distribution. For this reason, Nishimura and Ozaki (2007) and in a similar way Kort and Trojanowska (2010), considered a monopolistic firm which is uncertain about the distribution to use in order to evaluate the profit flow. Instead of using a single probability measure, they considered a set of measures that supports a set of possible drifts of the underlying stochastic process. Such a form of ambiguity is called drift ambiguity. Nishimura and Ozaki (2007) and Kort and Trojanowska (2010) assumed the monopolistic firm to be ambiguity–averse applying the max–min expected utility approach of Gilboa and Schmeidler (1989). The uncertainty about the drift is constructed using the so–called kappa–ignorance developed by Chen and Epstein (2002). Using kappa–ignorance, a firm considers all possible drifts in an interval $[\underline{\mu}, \overline{\mu}]$. Thijssen (2011) modeled the ambiguity not on the profit flow, but on the appropriate factor at which the cash flows are discounted. Even though the source of ambiguity was different, he came to the same result as Nishimura and Ozaki (2007) and Kort and Trojanowska (2010); an increase in ambiguity delays the investment of the firm. In such monopolistic models, the prior that yields the minimal expected utility is always given by the worst possible trend $\underline{\mu}$.

The inclusion of ambiguity, however, into a real options game has not been studied yet. This is what we do in chapter 4 of the present thesis.

We consider a preemption game with two firms being (potentially) heterogeneous in terms of both costs of investment and degree of uncertainty. The potentially low–cost firm is assumed to be uncertain about the drift of the underlying stochastic process using kappa–ignorance. The potentially high–cost firm, however, uses just a single prior. The fact that only one firm is ambiguous is, however, not important for our analysis. We model this in order to be able to indicate the effect ambiguity has on the payoff functions and equilibrium outcomes in contrast to a purely risky approach (see section 4.6 for a more detailed discussion on this point).

The main finding is that the determination of the payoff function and the worst–case scenario is no longer trivial under competition. We show that two opposing effects are at work; one supports the worst possible trend to be the right candidate for the worst–case, whereas the other one indicates exactly the opposite, namely that the worst–case is given by the best possible trend. Our main result demonstrates that the first effect always dominates the latter if the drop in the payoff due to competition is sufficiently small. In the case that this condition is not satisfied, we show that for the leader value function, there exists a unique threshold x^* such that the worst case is given by the lowest possible trend whenever the underlying process lies below x^*, and is given by the best possible trend if the underlying process lies above that threshold.

For the equilibrium concept, we use the framework developed in Riedel and Steg (2014). They extended the famous subgame perfect equilibrium concept given in Fudenberg and Tirole (1985) for deterministic timing games to stochastic timing games. To be able to include an ambiguous player, we just need to make some minor adjustments. The resulting equilibria can be either of preemptive or sequential type. We call an equilibrium a preemptive equilibrium whenever one of the firms is forced to invest in equilibrium sooner that it would do without the fear of competition (i.e. the fear of being preempted by its opponent). In a sequential equilibrium, such a fear does not exist and the firms invest at their optimal leader and follower threshold, respectively.

In contrast to the work of Pawlina and Kort (2006), which considered cost–asymmetric firms without ambiguity, we show that the high cost–firm may also become the leader in such a preemption game even in a sequential equilibrium.

This is the case if the cost disadvantage is sufficiently small compared to the degree of ambiguity.

Chapter 2

The Foster-Hart Measure of Riskiness for General Gambles[1]

2.1 Introduction

Risk measures are often used in the financial industry to determine a minimal capital reserve a company needs to have in order to overcome possible losses and ensure in that way its financial survival. In their famous work Artzner et al. (1999) introduced the class of coherent risk measures by stating four desirable axioms, namely monotonicity, homogeneity, subadditivity and cash invariance. Weakening the axioms homogeneity and subadditivity and replacing them by convexity, Frittelli and Rosazza Gianin (2002) and Föllmer and Schied (2002) proposed the more general class of convex risk measures.

These classes, however, are far away from defining a specific applicable measure. For this reason, Aumann and Serrano (2008) developed a specific index of riskiness that assigns to a random variable with known distribution a riskiness number. The Aumann–Serrano index of riskiness is defined as the reciprocal of the constant absolute risk aversion of an agent. Searching for an operational interpretation for this index, [2] Foster and Hart (2009) introduced

[1]Parts of this chapter were published in Riedel and Hellmann (2015).

[2]Meanwhile, Homm and Pigorsch (2012) found an operational interpretation even for the Aumann–Serrano index, showing its close connection to the adjustment coefficient which is a parameter for a ruin probability.

a new notion of riskiness, or critical wealth level, for random variables with known distribution. Their concept admits a simple operational interpretation because an agent avoids bankruptcy in the long run almost surely provided she accepts a gamble only if her current wealth exceeds the critical value. This measure is objective in so far as it depends only on the distribution of the outcome; in decision–theoretic terms, it is probabilistically sophisticated, in the language of the finance literature on risk measures, it is law–invariant.

Formally, the Foster–Hart measure of riskiness $R(X)$ is defined for a discrete random variable X on some probability space $(\Omega, \mathcal{F}, P)$ that satisfies $EX > 0$ and $P(X < 0) > 0$ by the unique solution of

$$E \log \left(1 + \frac{X}{R(X)}\right) = 0\,. \tag{2.1}$$

Foster and Hart (2009) argued that this number describes a threshold which divides between two different regimes: staying solvent with probability one on the one hand and the possibility of going bankrupt on the other. Consider, for instance, a gamble that pays with half of the probability \$200 and with half of the probability \$-100. We easily compute that the unique solution of equation (2.1) gives $R = 200$. The main theorem of Foster and Hart (2009) states now that a decision maker does not go bankrupt in the long run if and only if she rejects such gambles whenever her current wealth level is below \$200.

Foster and Hart (2009) and Hart (2011) presented a duality between the Aumann–Serrano index and the Foster–Hart measure of riskiness; the index is independent of the wealth of a decision maker and is based on the risk aversion whereas the riskiness measure focuses on the wealth level regardless of risk aversion. Furthermore, two new complete orders on gambles were introduced by Hart (2011), the wealth–uniform dominance and the utility–uniform dominance, showing that these orders are equivalent to the Aumann–Serrano index and the Foster–Hart measure of riskiness, respectively. Bali et al. (2011) provided a generalized version of both the Foster–Hart measure of riskiness and the Aumann–Serrano index of riskiness which involves a dependence on the risk aversion as well as on the wealth level of a decision-maker and finally both measures have been extended from gambles to securities [3] by Schreiber (2012).

Neither the Aumann–Serrano index nor the Foster-Hart measure belongs to the class of coherent risk measures. In fact, both fail only the cash invariance condition. However, Foster and Hart (2013) introduced another class of risk measures satisfying four basic axioms: dependence, homogeneity, monotonicity and compound gamble and showed that the minimal measure satisfying these four axioms is exactly the Foster–Hart measure.

Until now, the Foster-Hart measure of riskiness has only been studied for gambles with finitely many outcomes; even the finite examples were mostly confined to gambles with few values. Many financial applications involve distributions with a large number or a continuum of outcomes; it seems natural and important to generalize the concept of critical wealth level to such cases.

To our surprise, we realized that even for the most simple case of a uniform distribution, the defining equation of Foster and Hart does not always have a finite solution. The non–finite value of infinity is always a possible solution for the defining equation, but it would seem most counterintuitive and implausible to reject a uniformly distributed gamble on, say, the interval $[-100, 200]$ at arbitrary wealth levels. We show later that even for arbitrarily high gains greater than a certain critical value, the only solution for the defining equation is infinity. Clearly, there can come no useful theory out of a measure that suggests, regardless of how much money one may possess, to reject an uniform distributed gamble over, say, $[-100, 3.7 \times 10^{12}]$.

In this chapter we therefore set out to study the concept of riskiness for distributions with densities. We show that there are two classes of gambles. For some of them, the defining equation of Foster and Hart has a finite solution, and one can use this number as its riskiness. For others, the defining equation has no solution. We show that in this case, a reasonable extension is to use the *maximal loss* of the gamble as its riskiness. This might seem surprising at first sight, as for finite gambles, the Foster–Hart riskiness is always strictly

[3]The difference between gambles and securities lies in the fact on how they affect the wealth. If W_0 denotes the initial wealth level, then accepting a gamble g leads to a wealth of $W = W_0 + g$ and accepting a security r causes the wealth of $W = W_0 r$.

larger than the maximal loss. But we argue that it is the only rational way to extend the concept to arbitrary gambles.

As for discrete gambles the riskiness number is well defined, we approach this problem by studying an increasing sequence of discrete gambles (X_n) that converge to a continuous gamble X. We then investigate the asymptotic behavior of the respective riskiness numbers. For the uniform distribution above, an approximation involves a sequence of uniformly distributed finite gambles on a grid. As the grid becomes finer and finer, we show that the riskiness values converge to the maximal loss. This result carries over to gambles with arbitrary distributions. For gambles with a density where the Foster–Hart equation does have a solution, we also show that the riskiness values converge to that value. This is important as it provides a justification for using finite gambles with many outcomes as an approximation to non–finite gambles with a density.

A couple of examples eventually conclude this chapter.

2.2 Foster–Hart Model

For the sake of a better understanding of the Foster–Hart measure of riskiness, we briefly summarize the model proposed in Foster and Hart (2009).

In the Foster–Hart model, a decision maker with positive initial wealth level $W_0 > 0$ is offered a gamble X_t in every discrete time period $t \in \mathbb{N}$ that she can either accept or reject. A gamble X is a real valued discrete random variable, for which it holds that $P(X < 0) > 0$ and $E[X] > 0$. Accepting a gamble in period t leads to a wealth in the next period of $W_{t+1} = W_t + X_t$, whereas rejecting the gamble leaves the wealth unchanged, i.e. $W_{t+1} = W_t$. Further, let G denote the process of gambles $(X_t)_{t=1,2,\ldots}$. G is assumed to be finitely generated, i.e. each offered gamble X_t belongs to the finitely generated $cone\mathcal{G}_0 = \{\lambda X : \lambda \geq 0 \ and \ X \in \mathcal{G}_0\}$, where $\mathcal{G}_0$ denotes a finite set of gambles.

The decision maker is assumed to have a homogeneous critical wealth function $Q(X)$ which assigns to each gamble X a nonnegative real number such that she accepts X if and only if $W \geq Q(X)$. Furthermore, the decision maker is not allowed to borrow any money, i.e. $W_t \geq 0$ for all $t \in \mathbb{N}$. Bankruptcy is then defined as $\lim_{t\to\infty} W_t = 0$. Note, due to the no–borrowing condition, once the decision maker's wealth is zero, it stays zero forever.

The main result of Foster and Hart (2009) now states that, under these assumptions, a decision maker avoids bankruptcy for sure if and only if she rejects all offered gambles whenever her wealth is smaller than the assigned risk R of these positions, where R is given by the unique positive solution of equation (2.1).

2.3 Motivating Example and Notations

In the following, we consider on some probability space $(\Omega, \mathcal{F}, P)$ the reciprocal $\lambda = \frac{1}{R}$ of the Foster–Hart measure of riskiness, i.e. λ is the unique positive solution of

$$E \log(1 + X\lambda) = 0\,. \tag{2.2}$$

Note that for discrete random variables this equation is defined for all nonnegative values of λ up to, but strictly smaller than $\lambda^*(X) = 1/L(X)$, where $L(X) = \max_{\omega\in\Omega}(-X(\omega))$ is the maximal loss of the gamble.

For discrete random variables with positive expectation and possible losses, such a strictly positive solution always exists. For example, if X is a Bernoulli random variable with

$$P(X = 200) = P(X = -100) = \frac{1}{2}\,,$$

one can easily verify that $0 = \frac{1}{2}\log(1 + 200\lambda) + \frac{1}{2}\log(1 - 100\lambda)$ leads to $\lambda = 1/200$ or $R(X) = 200$.

The starting point of our analysis is the following simple observation that struck us when we wanted to apply the measure of riskiness to more general, continuous distributions.

Example 1 *Let X be uniformly distributed over $[-100, 200]$. X has the positive expectation 50 and losses occur with positive probability. It thus qualifies*

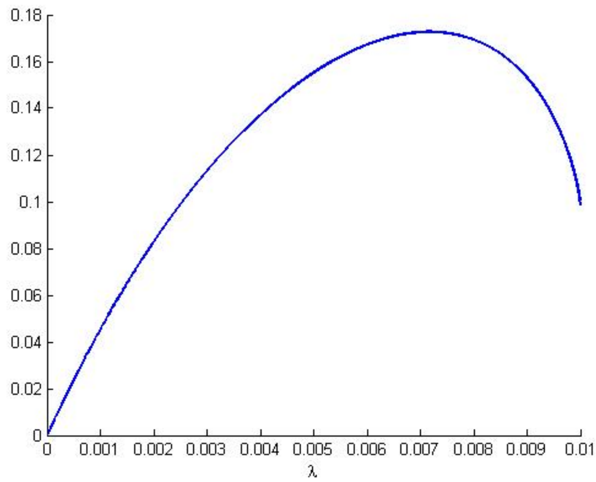


Figure 2.1: The function $\lambda \mapsto E \log(1 + \lambda X)$ for the uniform distribution over $[-100, 200]$ has no zero.

as a gamble in the sense of Foster and Hart. We study the equation

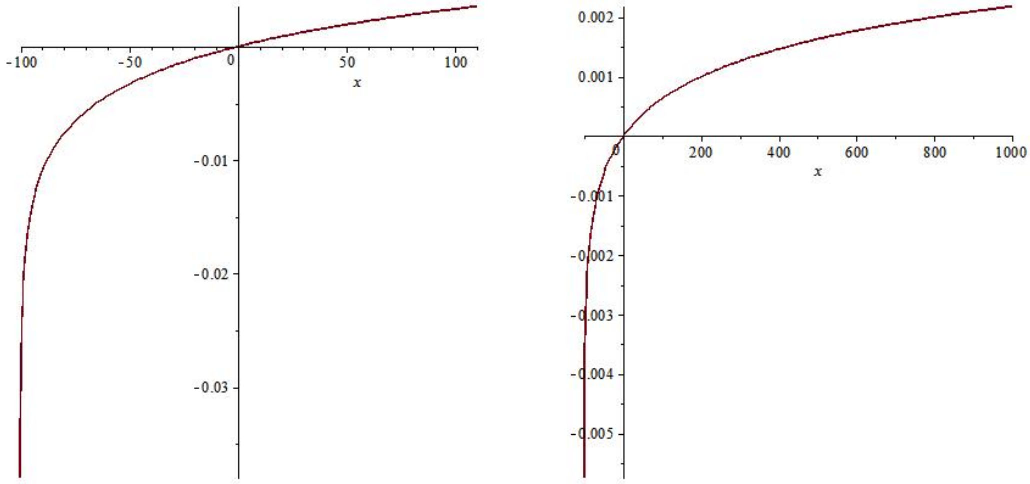
$$\phi(\lambda) := E \log(1 + \lambda X) = 0\,. \tag{2.3}$$

ϕ is well–defined for positive values $\lambda \leq 1/L$ where $L = 100$ is the gamble's maximal loss. We plot the function $\phi(\lambda)$ in Figure 2.1. No solution for $\lambda > 0$ to equation (2.3) *exists.*

For a formal proof, note that ϕ is continuous and concave on $[0, \frac{1}{L}]$, with positive slope in 0 as $EX > 0$ (see the argument in Foster and Hart (2009)). Thus, there exists a root for the defining equation if and only if $\phi(1/L) < 0$. For the maximal possible value $\lambda^*(X) = 1/L = 1/100$ we have

$$\begin{aligned} E \log(1 + \lambda^*(X) X) &= \int_{-100}^{200} \frac{1}{300} \log\left(1 + \frac{x}{100}\right) dx \\ &= \left[\frac{1}{3}\left(\left(1 + \frac{x}{100}\right) \log\left(1 + \frac{x}{100}\right) - \left(1 + \frac{x}{100}\right)\right)\right]_{-100}^{200} \\ &= \log 3 - 1 \simeq 0.0986 > 0\,. \end{aligned}$$

We conclude $\phi(\lambda) > 0$ for all $\lambda \in (0, \lambda^*(X)]$.



(a) $\frac{1}{210}\log\left(1+\frac{x}{100}\right)$ (b) $\frac{1}{1100}\log\left(1+\frac{x}{100}\right)$

Figure 2.2: The integrand of $E\log\left(1+\frac{X}{L}\right)$ for the uniform distributed gamble over $[-100, 110]$ and $[-100, 1000]$.

How can we explain the fact that no solution to the defining equation for such continuous gambles as the one above exists, whereas for discrete gambles always a solution can be obtained? To answer this question, consider again the function $\phi(\lambda) = E\log(1+\lambda X)$. We can easily verify that ϕ is continuous and concave on $[0, \frac{1}{L})$,[4] with positive slope in 0. Thus, there exists a root for the defining equation if and only if $\lim_{\lambda\to\frac{1}{L}}\phi(\lambda) < 0$. If X is discrete, its distribution places a strictly positive weight on the event $\{X = -L(X)\}$, where X achieves the maximal possible loss. As a consequence, the expression $\log(1-\lambda L(X))$ tends to minus infinity as λ approaches the value $\lambda^*(X) = 1/L(X)$. The expectation that defines ϕ then also converges to negative infinity for $\lambda \to \lambda^*(X)$, and the function ϕ has a unique zero in $(0, \lambda^*(X))$.

On the other hand, for continuous random variables X, the event $\{X = -L(X)\}$ has probability zero. In some cases, we have we have $\lim_{\lambda\to\frac{1}{L}}\phi(\lambda) = \phi(\frac{1}{L}) > 0$. In these cases, the distribution puts more weight on the positive area of the integral that defines the expectation than on the negative. Therefore,

[4]If X is continuous, the maximal loss L is not reached with a positive probability. Therefore, in contrast to the discrete case, equation (2.2) is also defined for $\lambda = \frac{1}{L}$ and ϕ is continuous and concave on $[0, \frac{1}{L}]$.

the expectation is positive for any $\lambda \in (0, \frac{1}{L}]$.

To illustrate that, we draw the integrand of $E \log \left(1 + \frac{X}{L}\right)$ for the uniform distributed gamble X with support $[-100, 110]$ in Figure 2.2(a) and with support $[-100, 1000]$ in Figure 2.2(b). In the first case, we observe that the negative area of the integral is greater than the positive. Therefore, the integral is negative which implies that a solution to the defining equation exists. In fact, one can show that $\lambda = 0.0027$ solves the defining equation. In the second case, more weight is put on the gain side of the distribution and the integral becomes positive. Hence, if the maximal loss is fixed, from a certain value on, there never exists a solution to the defining equation.

How can we assign a riskiness to a gamble for which the defining equation of Foster and Hart has no solution? One could take $\lambda = 0$, of course, resulting in a riskiness measure of ∞. Does this mean that one should never accept uniformly distributed gambles? Then an investor with a wealth of, say, a billion dollars would reject the above uniform gamble on $[-100, 200]$. Given that such a gamble has an expected gain of 50 and a maximal loss of 100, this would seem quite counterintuitive. Following the above argument, it would become even more counterintuitive if we consider the above uniform gamble with a much higher right bound of, say, one billion.

In this note, we therefore set out to extend the notion of riskiness for continuous random variables like the uniform above by approximating them via discrete random variables. We show that the limit of the riskiness coefficients exists. If the expectation $E \log(1 + \lambda^*(X)X)$ is negative (including negative infinity), one can use equation (2.2) to define the riskiness of X. (This also shows that our notion is the continuous extension of the discrete approach). For continuous random variables with $E \log(1 + \lambda^*(X)X) \geq 0$ such as our uniform random variable above, we use the limit of the riskiness coefficient of the approximating discrete random variables. This limit turns out to be equal to the maximal loss $L(X)$.

Whereas the riskiness measure is quite conservative for Bernoulli random variables as it prescribes a high value of 200 for the wealth for a Bernoulli random variable with maximal loss of 100, it does accept the uniform random variable over $[-100, 200]$, which has the same maximal loss of 100, even if one has just \$100.

How can one explain this difference? The point is that the Bernoulli random variable above is quite far away from the uniform random variable over the whole interval $[-100, 200]$. For the Bernoulli case, a loss of \$100 has a high probability of 50 %. For the uniformly distributed random variable, a loss of close to \$100 occurs with a very small probability and the loss of exactly \$100 even with probability zero; as the defining aim of the operational measure of riskiness is to avoid bankruptcy, this is a crucial difference. Indeed, our analysis below shows that the riskiness decreases if we spread the Bernoulli random variable more uniformly over the interval $[-100, 200]$, say by using a uniform grid. For discrete random variables uniformly distributed over a fine grid, the riskiness is close to the maximal loss of 100.

Let us study next how the riskiness numbers of discrete distributions that approximate the uniform one look like. We approximate the uniform distribution over $[-100, 200]$ by finite gambles.

Example 2 *We consider discrete and uniformly distributed gambles on the grid* $-100, -100 + \frac{300}{n-1}, \ldots, -100 + \frac{300k}{n-1}, \ldots, 200$. *The riskiness is the root of*

$$f_n(\lambda) = \frac{1}{n} \sum_{k=0}^{n-1} \log\left(1 + \lambda\left(-100 + \frac{300k}{n-1}\right)\right).$$

For the simplest case $n = 2$ *we showed above that* $R = \frac{1}{\lambda} = 200$.

In the following table the riskiness numbers for different grid sizes are shown. We observe that the riskiness number decrease and converge to the maximal loss as the grid becomes finer and finer. As the single weights on specific losses vanish, the investor might accept the gambles at ever lower wealth levels. In the limit, she is able to gamble as long as her wealth suffices to cover the maximal loss without taking any risk of bankruptcy.

n	*grid size*	*riskiness*
2	*300*	*200*
3	*150*	*145.74*
5	*75*	*119.46*
11	*30*	*104.997*
21	*15*	*101.197*
31	*10*	*100.3651*
41	*7.5*	*100.1195*
61	*5*	*100.0137*
101	*3*	*100.0002*

2.4 Main Result

We now go beyond specific examples and clarify for which gambles the Foster–Hart index is equal or close to the maximal loss. We will characterize such distributions by a simple condition.

Before we start, let us fix what we mean by gamble.

Definition 2 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. We call a random variable $X : (\Omega, \mathcal{F}) \to \mathbb{R}$ a gamble if*

- *its expectation is positive: $EX > 0$,*
- *losses occur with positive probability: $P[X < 0] > 0$,*
- *and its maximal loss is bounded: $L(X) := \operatorname{ess\,sup}(-X) < \infty$* [5]*.*

We call a gamble finite if its support is finite.

Let us now describe how one could construct an approximating sequence of discrete gambles more formally. We approximate X from below by an increasing sequence of discrete random variables $X_n \uparrow X$. For each X_n, there is a unique positive number $0 < \lambda_n < \lambda^*(X)$ that solves the defining equation

$$E \log (1 + \lambda_n X_n) = 0 \,.$$

[5] We define as usual $\operatorname{ess\,sup}(-X) := \inf\{x \in \mathbb{R} | P(-X > x) = 0\}$.

The sequence (λ_n) is increasing and bounded, thus the limit $\lambda_\infty = \lim_{n\to\infty} \lambda_n$ exists and is bounded by $\lambda^*(X)$. We will show that for random variables X with $\phi(\lambda^*(X)) = E \log (1 + \lambda^*(X)X) < 0$, λ_∞ is the unique positive solution of the defining equation (2.2). For random variables like the uniform one, where $\phi(\lambda^*(X)) = E \log (1 + \lambda^*(X)X) \geq 0$, we have $\lambda_\infty = \lambda^*(X)$.

Without loss of generality, we take $L = 1$ and thus $\lambda^*(X) = 1$. We consider two different sequences of partitions of the support of the continuous random variables.

On the one hand, if the support of X is the compact interval $[-1, M]$, we define

$$x_k^n = -1 + \frac{k}{2^n}(M+1), \qquad k = 0, \ldots, 2^n - 1$$

and set

$$X_n = \sum_{k=0}^{2^n-1} x_k^n 1_{\left\{x_k^n \leq X < x_{k+1}^n\right\}}.$$

On the other hand, if the support of X is the infinite interval $[-1, \infty)$, we define

$$x_k^n = -1 + \frac{k}{2^n}, \qquad k = 0, \ldots, n2^n - 1$$

and set for $n \geq 1$

$$X_n = \sum_{k=0}^{n2^n-1} x_k^n 1_{\left\{x_k^n \leq X < x_{k+1}^n\right\}} + (-1+n) 1_{\{X \geq (-1+n)\}}.$$

For both cases the next Lemma holds true.

Lemma 1 *The sequence* (X_n) *is increasing and* $\lim X_n = X$ *a.s.*

Proof: Consider first the case where the support of X is given by a compact interval $[-L, M]$. Obviously, we have $x_k^n < x_{k+1}^n$.

Further,

$$x_k^n = -1 + \frac{k}{2^n}(M+1) = -1 + \frac{2k}{2^{n+1}}(M+1) = x_{2k}^{n+1}.$$

Therefore, we get

$$X_n = \sum_{k=0}^{2^n-1} x_k^n 1_{\left\{x_k^n \le X < x_{k+1}^n\right\}} = \sum_{k=0}^{2^n-1} x_{2k}^{n+1} 1_{\left\{x_{2k}^{n+1} \le X < x_{2(k+1)}^{n+1}\right\}}.$$

The last expression is smaller than X_{n+1} since we have on each interval $[x_{2k}^{n+1}, x_{2(k+1)}^{n+1}) \neq \emptyset$ that

$$X_n = x_{2k}^{n+1}$$

and

$$X_{n+1} \geq x_{2k}^{n+1}$$

with $X_{n+1} > x_{2k}^{n+1}$ on $\emptyset \neq [x_{2k+1}^{n+1}, x_{2(k+1)}^{n+1}) \subset [x_{2k}^{n+1}, x_{2(k+1)}^{n+1})$. Hence, (X_n) is an increasing sequence.

Analogously, we can prove the statement for the support $[-1, \infty)$, just by setting x_k^n equal to $-1 + \frac{k}{2^n}$ and X_n to

$$\sum_{k=0}^{n2^n-1} x_k^n 1_{\left\{x_k^n \le X < x_{k+1}^n\right\}} + (-1+n) 1_{\{X \geq (-1+m)\}}.$$

Finally, we have $\lim X_n = X$ $a.s.$ by construction of the sequence (X_n). □

As we have $-1 \leq X_n \leq X \in L^1$,[6] we conclude, using Lebesgue dominated convergence theorem, that $\lim EX_n = E \lim X_n = EX > 0$. Hence, for n sufficiently large, we have $EX_n > 0$. From now on, we always look at such large n only. As the density of X is strictly positive on its support, we also have $P(X_n < 0) \geq P(X_n = -1) > 0$. Therefore, the Foster–Hart riskiness is well–defined for X_n. Let $\lambda_n \in (0, 1)$ be the unique positive solution of

$$E \log(1 + \lambda_n X_n) = 0\,.$$

The next Lemma follows directly by Lemma 1 and by the monotonicity of the Foster–Hart measure of riskiness, see Proposition 2 in Foster and Hart (2009).

[6]We denote by L^1 the space of all random variables X with $EX < \infty$.

Lemma 2 *The sequence (λ_n) is increasing and bounded by $L(X) = 1$. As a consequence,*

$$\lambda_\infty = \lim \lambda_n$$

exists and is less or equal to $L(X) = 1$.

We can now state our main theorem. The limit λ_∞ identified in the previous lemma is the right tool to define the riskiness for general gambles.

Theorem 1 *Let X be a gamble with maximal loss $L > 0$. Let (X_n) be a sequence of finite gambles with $X_n \uparrow X$ a.s., where each X_n has the same maximal loss L. Denote by $\rho_n := \rho(X_n) > L$ their Foster–Hart riskiness. Then the following holds true:*

1. *If $E \log(1 + X/L) < 0$, then $\rho_\infty > L$ and ρ_∞ is the unique positive solution of the Foster–Hart equation (2.1).*
2. *If $E \log(1 + X/L) \geq 0$, then the Foster–Hart equation has no solution and $\rho_\infty = L(X)$.*

PROOF:

It is easier to prove the converse of the two statements. Without loss of generality, we take $L = 1$ (else replace X by X/L). Let us start with assuming $\lambda_\infty < 1$. In that case, the sequence

$$Z_n = \log\left(1 + \lambda_n X_n\right)$$

is uniformly bounded. Indeed,

$$-\infty < \log(1 - \lambda_\infty) \leq Z_n \leq \log(1 + |X|) \leq |X| \in L^1 .$$

As we have $Z_n \to \log(1 + \lambda_\infty X)$ a.s., we can then invoke Lebesgue's dominated convergence theorem to conclude

$$0 = \lim E Z_n = E \lim Z_n = E \log(1 + \lambda_\infty X) .$$

In particular, the equation (2.3) has a positive solution $\lambda_\infty < 1$. As $\phi(\lambda) = E\log(1+\lambda X)$ is strictly concave and strictly positive on $(0, \lambda_\infty)$, we conclude that we must have $\phi(1) = E\log(1+X) < 0$. This proves the second claim.

Now let us assume $\lambda_\infty = 1$. In that case, we cannot use Lebesgue's theorem. However, the sequence

$$Z'_n = -\log\left(1 + \lambda_n X_n\right)$$

is bounded from below by $-\log(1+|X|) \geq -|X| \in L^1$. We can then apply Fatou's lemma to conclude

$$-E\log(1+X) = E\lim Z'_n \leq \liminf -E\log\left(1+\lambda_n X_n\right) = 0\,,$$

or

$$E\log(1+X) \geq 0\,.$$

This proves the first claim. □

After stating our main theorem, we are eventually able to define our extended Foster–Hart measure of riskiness.

Definition 3 *Let X be a gamble. If $E\log(1+X/L) < 0$, we define the extended Foster–Hart measure of riskiness $\rho(X)$ as the unique positive solution of equation*

$$E\log\left(1 + \frac{X}{\rho(X)}\right) = 0\,.$$

If $E\log(1+X/L) \geq 0$, we define $\rho(X)$ as the maximal loss of X,

$$\rho(X) = L\,.$$

Remark 1 *Foster and Hart (2009) noticed that the measure of riskiness is not necessarily continuous, meaning that for a sequence of gambles (X_n) which converges in distribution to a gamble X, the limit of the respective riskiness numbers is not necessarily $\rho(X)$. This, however, can only be the case if the sequence of maximal losses $(L(X_n))$ does not converge to $L(X)$. It is therefore important that we fix the maximal loss of each gamble of the approximating sequence (X_n) equal to the maximal loss of X.*

Further note, for the construction of the sequence (X_n) *the increase with respect to first order stochastic dominance of* (X_n) *is important as it has a clear interpretation; higher gains and/or lower losses strictly decrease* [7] *the riskiness number. Thus, since* $X_n \leq X$ *we have* $\rho(X_n) \geq \rho(X) \geq L$ *and we showed that in the limit* $\rho(X_n) \to \rho(X)$.

For gambles where the defining equation does not have a solution, our theorem suggests to use the maximal loss as their riskiness.

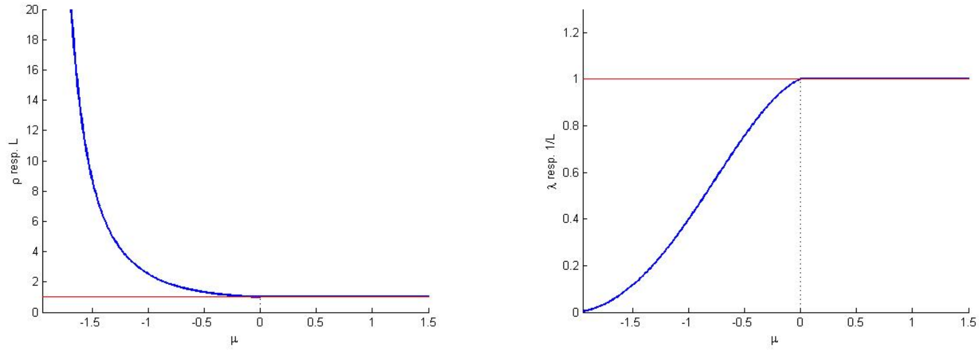
The previous theorem also gives a simple test to see whether the Foster–Hart riskiness is equal (or close to) the maximal loss of a distribution. Indeed, the sign of the expectation $E \log(1 + X/L)$ determines whether the riskiness is equal or close to the maximal loss.

The maximal loss is indeed obtained for a large number of gambles. For example, for the uniform distribution on $[-100, 200]$ and for a uniform distribution on, say, $[-100, 10^{12}]$, the riskiness is the same, namely 100 (and similarly for finite gambles with such a support on a dense grid, compare Example (2) above for more details). The Foster–Hart riskiness index then boils down to the so–called worst–case risk measure.

This property appears to be undesirable. Why would uniform gambles on $[-100, 200]$, and the much more favorable uniform gambles on $[-100, 10^{12}]$ have the same riskiness? Let us look at the operational interpretation of the riskiness that Foster and Hart had in mind. The aim is to find a critical wealth level that ensures solvency with probability one if it is used as a decision rule for acceptance and rejection of gambles. For solvency, losses clearly play a much more important role than potential gains (and of course one needs to have at least the maximal loss in order to guarantee no–bankruptcy even for a favorable gamble like the uniform on $[-100, 10^{12}]$). Our analysis shows that frequently the maximal loss only determines whether one should accept or reject a gamble.

In the next chapter, we extend the Foster–Hart result on solvency to our gambles. A decision maker does avoid bankruptcy with probability one if she

[7] In case of continuous distributed gambles, the sequence of Foster–Hart riskiness numbers of a monotone increasing, with respect to first order stochastic dominance, sequence of gambles is at least monotone decreasing but not necessarily strictly anymore.



(a) Generalized Foster–Hart riskiness ρ (b) $\lambda = \frac{1}{\rho}$

Figure 2.3: ρ and λ for lognormally distributed gambles over $(-1, \infty)$ with $\sigma = 2$.

uses our extended riskiness as a decision rule. In particular, if a decision maker faces a sequence of independent uniformly distributed gambles with sufficiently high maximal gains, she stays solvent with probability one if she accepts every gamble which maximal loss is below her wealth. As the operational interpretation of Foster and Hart (2009) carries over, this provides another justification for using the maximal loss as an extension of the Foster–Hart riskiness.

Our theorem also shows that, for certain gambles, the Foster–Hart index does not care about the way gains are distributed. Whether you have specific gains with a certain density or point masses on some numbers does not matter. Further examples given in the next section illustrate this point.

2.5 A list of examples

2.5.1 Lognormal Distribution

The lognormal distribution is used in many financial applications, for instance in the widely used Black–Scholes options pricing model. It seems, therefore, important to be able to apply the measure of riskiness for this distribution.

A random variable X is said to be lognormally distributed if its density φ

is

$$\varphi(x;\mu,\sigma,L) = \frac{1}{(x+L)\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(log(x+L)-\mu)^2}{\sigma^2}\right), \quad x > -L,^8$$

where μ and σ are the expected value and the standard deviation, respectively, of the normally distributed random variable $X^N = \log(X+L)$ and L is the maximal loss of X.

For the special case of lognormally distributed gambles with $L = 1$, we can obtain an interesting result.

Proposition 1 *For the lognormally distributed random variable* $X = \exp(X^N) - 1$ *with* $EX > 0$*, there exists a solution for the defining equation* (2.1) *if and only if* $EX^N < 0$*.*

Proof: We can easily check that

$$E\log\left(1 + \frac{\exp(X^N) - 1}{1}\right) = E\log(\exp(X^N)) = EX^N$$

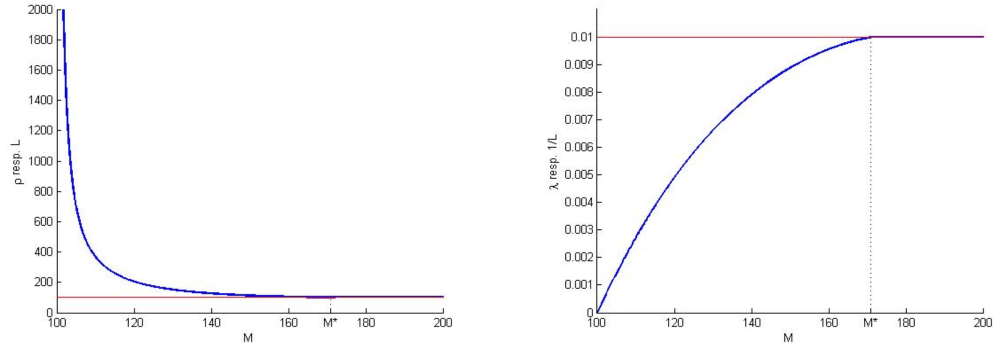
and therefore $E\log(1 + \frac{X}{L(X)}) < 0$ if and only if $EX^N < 0$. □

Now, if we also fix $\sigma = 2$, we can numerically compute the riskiness as a function of μ. The result is drawn in Figure 2.3. As Proposition 1 already says, we observe that the critical value for which there exists no zero for the defining equation is $\mu^* = 0$.

2.5.2 Uniform Distribution

Let us consider the motivating example again. We fix $L = 100$ and check for which value M^* of the maximal gain the defining equation (2.1) has a solution for the uniformly distributed gamble over $[-100, M^*]$, i.e. we need to find M^*

[8] See Johnson et al. (1995).



(a) Generalized Foster–Hart riskiness ρ (b) $\lambda = \frac{1}{\rho}$

Figure 2.4: ρ and λ for uniformly distributed gambles over $[-100, M]$.

such that $E \log\left(1 + \frac{X}{L}\right) = 0$. Therefore,

$$
\begin{aligned}
E \log\left(1 + \frac{X}{100}\right) &= \int_{-100}^{M^*} \frac{1}{100 + M^*} \log\left(1 + \frac{x}{100}\right) dx \\
&= \left[\frac{100}{100 + M^*}\left(\left(1 + \frac{x}{100}\right)\log\left(1 + \frac{x}{100}\right) - \left(1 + \frac{x}{100}\right)\right)\right]_{-100}^{M^*} \\
&= \frac{100}{100 + M^*}\left(\left(1 + \frac{M^*}{100}\right)\log\left(1 + \frac{M^*}{100}\right) - \left(1 + \frac{M^*}{100}\right)\right).
\end{aligned}
$$

Setting this equal to zero yields

$$\log\left(1 + \frac{M^*}{100}\right) = 1,$$

which implies

$$M^* = L(e-1) \simeq 171.8.$$

Hence, for all values $M < M^*$ there exists a solution to the defining equation and we take this solution as the riskiness. For all $M \geq M^*$ there does not exists a finite solution and we take therefore the maximal loss $L = 100$ as the riskiness.

In Figure 2.4 the graph of the riskiness ρ as well as the solution λ of equation (2.3) is plotted against the maximal gain M of the gambles. In order to have a positive expectation, we consider only values of M with $M > 100$.

As a result the graph of the riskiness shows a continuous function; the riskiness tends to the maximal loss $L = 100$ as we approach the critical value M^* and converges to infinity as the expectation of the gamble goes to 0 (i.e. $M \downarrow 100$).

2.5.3 Mixed Distribution

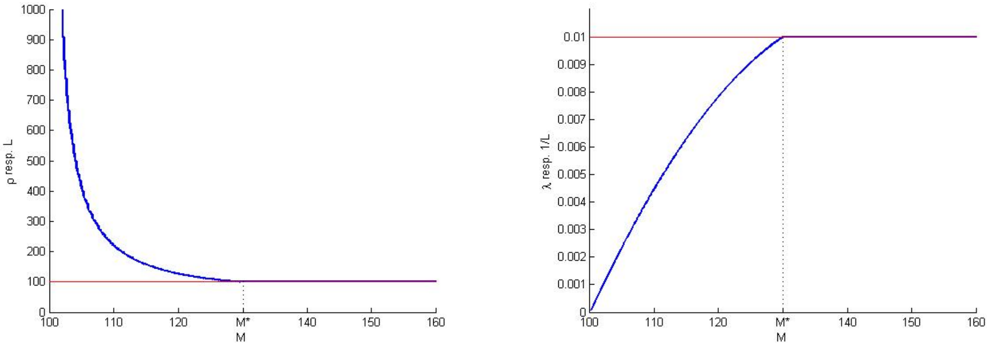
Consider mixed gambles that have a discrete part as well as a continuous one. For instance, we take a gamble X that is uniformly distributed over the interval $[-100, 0]$ and that places a probability of 50% on the event $\{X = M\}$, where $M > 50$ to ensure a positive expectation. For $M > M^* := 100(e-1) \simeq 171.8$, no solution to the defining equation exists as a similar calculation as in the previous example shows:

$$\begin{aligned}
E \log\left(1 + \frac{X}{100}\right) &= \frac{1}{2}\log\left(1 + \frac{M}{100}\right) + \frac{1}{2}\int_{-100}^{0} \frac{1}{100}\log\left(1 + \frac{x}{100}\right) dx \\
&= \frac{1}{2}\Bigg(\log\left(1 + \frac{M}{100}\right) + \left(1 + \frac{0}{100}\right)\log\left(1 + \frac{0}{100}\right) \\
&\qquad\qquad\qquad - \left(1 + \frac{0}{100}\right)\Bigg) \\
&= \frac{1}{2}\left(\log\left(1 + \frac{M}{100}\right) - 1\right).
\end{aligned}$$

For $M > M^*$, this expression is positive.

We observe that the critical value for this mixed distribution is the same as for the uniform distribution over $[-100, 171.8]$ (see Example (2.5.2)). This seems to be surprising at first sight as we replaced an uniform distribution over an interval by a positive mass on the maximal gain. But notice that the Foster–Hart measure of riskiness is more sensitive on the loss side than on the gain side. The decrease of the probability of the event $\{X \leq 0\}$ from $\frac{1}{2}$ to $\simeq 0.37$ vanishes the higher gains of the mixed gamble and the critical value is exactly the same.

On the other hand, if we take a mixed distribution that has a point mass



(a) Generalized Foster–Hart riskiness ρ (b) $\lambda = \frac{1}{\rho}$

Figure 2.5: ρ and λ for beta distributed gambles over $[-100, 200]$ with $\alpha = 2$.

on its maximal loss L, the defining equation always posses a solution. Indeed, due to the fact that the event $\{X = -L\}$ has a positive probability, we have $\lim_{\lambda \to \frac{1}{L}} E \log(1 + \lambda X) = -\infty$ and therefore a solution to equation (2.1) exists.

2.5.4 Beta Distribution

Let us consider beta distributed gambles. The density of a random variable X that is beta distributed over the compact interval $[-L, M]$ is, for instance, given in Johnson et al. (1995) as

$$\varphi(x; \alpha, \beta, L, M) = \frac{1}{B(\alpha, \beta)} \frac{(x + L)^{\alpha - 1}(M - x)^{\beta - 1}}{(M + L)^{\alpha + \beta - 1}}, x \in [-L, M], \alpha, \beta > 0,$$

where $B(\alpha, \beta)$ denotes the Betafunction defined as

$$B(\alpha, \beta) = \int_0^1 t^{\alpha - 1}(1 - t)^{\beta - 1} dt.$$

The mean of X is given by

$$EX = \frac{\alpha M - \beta L}{\alpha + \beta}.$$

We can parameterize our beta distributed gamble X by

$$X = cZ - L,$$

where Z is a beta distributed random variable over $[0, 1]$ and $c = M + L$. Using this parameterization, we can now explicitly compute for which value of M (or c) no solution to the defining equation exists. Let us fix $L = 100$, $\alpha = 2$ and $\beta = 2$. We have

$$\begin{aligned} E\log\left(1+\frac{X}{L}\right) &= E\log\left(1+\frac{cZ-L}{L}\right) \\ &= E\log\left(\frac{cZ}{L}\right) \\ &= \log(c) - \log(L) + E\log(Z). \end{aligned}$$

Thus, we are searching for c^* that solves

$$\log(c^*) = \log(L) - E\log(Z).$$

Now,

$$\begin{aligned} E\log(Z) &= \int_0^1 \frac{\log(x)}{B(2,2)} x(1-x)dx \\ &= \frac{1}{B(2,2)}\left[\log(x)\left(\frac{1}{2}x^2 - \frac{1}{3}x^3\right) - \left(\frac{1}{4}x^2 - \frac{1}{9}x^3\right)\right]_0^1 \\ &= -\frac{5}{6}. \end{aligned}$$

Hence,

$$c^* = 100\exp\left(\frac{5}{6}\right) \simeq 230.09$$

which means

$$M^* \simeq 130.09.$$

Figure 2.5 shows the graph of the riskiness and of $\lambda = \frac{1}{\rho}$ against M. For

$M > M^*$, where no positive solution exists, the maximal loss $L = 100$ is used to determine the riskiness. The figure again demonstrates that the maximal loss is a continuous extension of the Foster–Hart measure of riskiness.

Chapter 3

A Dynamic Extension of the Foster–Hart Measure of Riskiness[1]

3.1 Introduction

Foster and Hart (2009) introduced a notion of riskiness, or critical wealth level, for gambles with known distribution. Formally, the Foster–Hart measure of riskiness is given by the unique solution $R(X)$ of

$$E \log\left(1 + \frac{X}{R(X)}\right) = 0 \,. \tag{3.1}$$

The Foster–Hart measure of riskiness $R(X)$ is defined for discrete random variables X on some probability space $(\Omega, \mathcal{F}, P)$ that satisfy $EX > 0$ and $P(X < 0) > 0$.

In chapter 2 we noticed that for general continuous distributions the defining equation does not necessarily admit a solution. In this case, the riskiness numbers of sequences of discrete gambles that approximate the gamble with continuous distribution converge to the maximal loss of the gambles. We thus suggest to use the maximal loss as the reasonable extension for the Foster-Hart

[1] This chapter was published in Hellmann and Riedel (2015).

measure when there exists no solution to equation (3.1).

In this chapter, we study the extended Foster–Hart index of riskiness for general gambles in dynamic settings. As many financial applications require to quantify risk over time in a dynamic way, it seems natural and important to generalize the concept to a dynamic framework.

Dynamic measurement of risk plays an important role in the recent literature [2] since it allows, in contrast to the static case, to measure risk of financial positions over time. The arrival of new information can thus be taken into account. This is important for many situations; suppose, for instance, one faces a gamble that has its payments in, say, one month. In two weeks from now the information about this gamble might be much more precise which allows to adjust the risk assessment and to determine the risk more accurately. A static risk measure cannot do that. To cover such cases it is therefore crucial to be able to merge from static to dynamic risk measurement.

We thus set out to study the Foster–Hart measure of riskiness (or more precisely the extended Foster–Hart measure of riskiness defined in chapter 2) in a dynamic framework. As a first step, we define the concept of conditional Foster–Hart riskiness for general probability spaces and filtrations. In the original work of Foster and Hart (2009) a somewhat dynamic approach is already needed to prove the no–bankruptcy result. Their approach, however, is rather intuitive than precise in a measure–theoretic sense. We provide here a more rigorous approach which allows us also to drop the assumption used in Foster and Hart (2009) that all gambles are multiples of a finite number of so–called basic gambles.[3] Furthermore, we allow the extended Foster–Hart measure of riskiness to measure also gambles with potentially unbounded gains.

In the new framework, we show that Foster–Hart's no-bankruptcy result (and with it the operational interpretation) carries over to general continuous distributions. The proof uses a different martingale argument which might be interesting in itself.

A desirable property of a dynamic risk measure is the notion of time–

[2]See, among others, Detlefsen and Scandolo (2005) and Föllmer and Schied (2011), Chapter 11 for a detailed introduction to dynamic risk measures.

[3]See section 2.2 for a description of the assumptions used by Foster and Hart (2009).

consistency. Time–consistency for dynamic risk measures is widely studied in the recent literature, see, among others, Riedel (2004), Roorda et al. (2005), Detlefsen and Scandolo (2005), Weber (2006) and Artzner et al. (2007). Roughly speaking a measure is time–consistent if it assigns a greater risk to one gamble than to another whenever it is known that the same holds true tomorrow. This property yields a consistent behavior of an agent who bases her decision on a time–consistent risk measure.

This property is not satisfied by many risk measures. In fact, the still most widely used Value at Risk has, besides many other undesirable properties, this inconsistence feature as it is shown in Cheridito and Stadje (2009). The same holds true for the dynamic Average Value at Risk. Cheridito and Stadje (2009), however, propose an alternative time–consistent version of the Value at Risk by composing one period Value at Risks over time.

On the other hand, a nice example for a time–consistent risk measure is given in Detlefsen and Scandolo (2005). They show that the dynamic entropic risk measure which is closely related to an agent with expected exponential utility preferences is time–consistent.

The dynamic version of the Foster–Hart measure of riskiness, however, does not satisfy the time–consistency condition. We show this by the use of a simple two period example. This example indicates a difference between the original static Foster–Hart measure and our dynamic version. In some instances the static Foster–Hart measure differentiates between two gambles, which are assigned to the same risk in every possible state of the world at a certain point in time by the conditional measure.

The chapter is set up as follows: Section 2 introduces the dynamic framework as well as the dynamic extended Foster–Hart measure of riskiness. In Section 3 we give the more general no–bankruptcy result. Section 4 contains a counterexample which shows the time–inconsistency of the new defined measure. Finally, we prove the existence of the dynamic Foster–Hart index in Section 5.

3.2 The Dynamic Framework

In the following, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}}, P)$ be a filtered probability space. We denote by $\mathcal{A}_t$ the set of all $\mathcal{F}_t-$measurable random variables and consider a sequence of random variables (X_t) that is adapted to the filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$. In order to be able to measure the risk of X_t in every time period $s < t$, X_t has to satisfy all the conditions of Definition (2) given the filtration $(\mathcal{F}_s)$.

Definition 4 *We call a random variable $X \in L^2$ on $(\Omega, \mathcal{F}, P)$ a gamble for the σ–field $\mathcal{F}_s \subset \mathcal{F}$ if X is bounded from below and satisfies $E[X|\mathcal{F}_s] > 0$ a.s. and $P(X_t < 0|\mathcal{F}_s) > 0$ a.s.*

In the remainder, we assume that for $t > s$, X_t is a gamble for $\mathcal{F}_s$. We denote by $L_s(X_t)$ the maximal loss of X_t given the information at time s. Formally,

$$L_s(X_t) := \operatorname{ess\,inf}\{Z \in \mathcal{A}_s | P(-X_t > Z|\mathcal{F}_s) = 0 \ \ a.s.\}.$$

We now embed the extended riskiness notion of chapter 2 in the dynamic framework. As time goes by, we learn something about the realization of the random variable and are therefore able to quantify the risk more precisely. Measuring the risk of X_t in every single time period $s < t$ yields a family of conditional risk measures $(\rho_s(X_t))_{s=1\ldots t-1}$, where every $\rho_s(X_t)$ is a $\mathcal{F}_s-$measurable random variable. For continuous random variables the equation

$$E\left[\log\left(1 + \frac{X_t}{\rho_s(X_t)}\right)|\mathcal{F}_s\right] = 0 \tag{3.2}$$

does not always have a solution. Following the arguments given in chapter 2[4], this is the case on the set

$$\mathcal{B} := \left\{E\left[\log\left(1 + \frac{X_t}{L_s(X_t)}\right)|\mathcal{F}_s\right] \geq 0\right\}.$$

As in the static case, on $\mathcal{B}$ the conditional maximal loss is the reasonable extension of the classical riskiness notion.

[4]For more details we refer to Section 3.5.

The next theorem shows that what we define later as the dynamic extended Foster–Hart riskiness is well defined.

Theorem 2 *There exists one and only one $\mathcal{F}_s$–measurable random variable $\rho_s(X_t) \geq L_s(X_t)$ that solves equation* (3.2) *on $\mathcal{B}^c$ and satisfies $\rho_s(X_t) = L_s(X_t)$ on $\mathcal{B}$.*

We give the technical proof of the theorem in Section 3.5.

We are now ready to give the definition of the dynamic extended Foster–Hart measure of riskiness.

Definition 5 *The dynamic extended Foster–Hart measure of riskiness for a gamble X_t is the family of conditional risk measures $(\rho_s(X_t))_{s=1\ldots t-1}$, where each $\rho_s(X_t)$ is equal to the conditional maximal loss $L_s(X_t)$ on $\mathcal{B}$ and the solution to equation* (3.2) *on $\mathcal{B}^c$.*

3.3 No–Bankruptcy Result

The main result of Foster and Hart (2009) yields that a decision maker who rejects a gamble whenever her wealth is below the assigned riskiness number avoids bankruptcy (with probability one). It is crucial not to lose this property (and with it the operational interpretation of the measure) when working with continuous distributed gambles.

We provide here the respective no–bankruptcy theorem for the extended Foster–Hart measure of riskiness.

Theorem 3 *Let (X_n) be a sequence of gambles that are uniformly bounded above by some integrable random variable $Y > 0$ and satisfy some minimal possible loss requirement, i.e. there exists $\epsilon > 0$ such that a.s.*

$$L_{n-1}(X_n) \geq \epsilon > 0$$

for all n. Let $W_0 > 0$ be the initial wealth and define recursively

$$W_{t+1} = W_t + X_{t+1}$$

if $E\left[\log\left(1 + X_{t+1}/W_t\right)|\mathcal{F}_t\right] \geq 0$ *and*

$$W_{t+1} = W_t$$

else. We then ensure no–bankruptcy, i.e.

$$P[\lim W_t = 0] = 0\,.$$

PROOF: Throughout the proof, we assume that all inequalities and equalities between random variables hold $P-$almost surely.

Note first that $W_t > 0$. This can be shown by induction. We have $W_0 > 0$. We have either $W_{t+1} = W_t$ which is positive by induction hypothesis, or $W_{t+1} = W_t + X_{t+1}$. In this case, the condition $E\left[\log\left(1 + X_{t+1}/W_t\right)|\mathcal{F}_t\right] \geq 0$ implies that

$$W_t \geq \rho_t(X_{t+1}) \geq L_t(X_{t+1}).$$

Thus, $W_t - L_t(X_{t+1}) \geq 0$. The maximal loss can only be obtained by the riskiness measure if the considered gamble is continuous. Therefore, if $\rho_t(X_{t+1}) = L_t(X_{t+1})$, we have $P(X_{t+1} = L_t(X_{t+1})|\mathcal{F}_t) = 0$. Hence, it holds that

$$W_{t+1} > W_t - L_t(X_{t+1}) \geq 0.$$

We can thus define $S_t = \log W_t$. We claim that S is a submartingale. Indeed, on the set

$$A := \left\{E\left[\log\left(1 + \frac{X_{t+1}}{W_t}\right)|\mathcal{F}_t\right] < 0\right\}$$

which belongs to $\mathcal{F}_t$, there is nothing to show. On the set A^c, we have

$$\begin{aligned}
E\left[S_{t+1}|\mathcal{F}_t\right] &= E\left[\log W_{t+1}|\mathcal{F}_t\right] \\
&= \log W_t + E\left[\log\frac{W_{t+1}}{W_t}|\mathcal{F}_t\right] \\
&= \log W_t + E\left[\log\left(1 + \frac{X_{t+1}}{W_t}\right)|\mathcal{F}_t\right] \\
&\geq \log W_t = S_t\,.
\end{aligned}$$

S is thus a submartingale. We apply the theorem on submartingale convergence in Shiryaev (1984), Chapter VII, Theorem 1. For $a > 0$, let $\tau_a = \inf\{t \geq 0 : X_t > a\}$. A stochastic sequence belongs to class $\mathcal{C}^+$ if for every $a > 0$ we have

$$E\left(X_{\tau_a} - X_{\tau_a - 1}\right)^+ 1_{\{\tau_a < \infty\}} < \infty.$$

Let us check that our sequence S is of class $\mathcal{C}^+$. Indeed, we have

$$\left(S_{\tau_a} - S_{\tau_a - 1}\right)^+ = \log\left(1 + \frac{X_{\tau_a}}{W_{\tau_a - 1}}\right) 1_{\{X_{\tau_a} \geq 0\}}$$

and in that case

$$W_{\tau_a - 1} \geq \rho_{\tau_a - 1}(X_{\tau_a}) \geq \epsilon > 0.$$

Hence, we conclude

$$E\left(S_{\tau_a} - S_{\tau_a - 1}\right)^+ \leq E \log\left(1 + \frac{Y}{\rho_{\tau_a - 1}(X_{\tau_a})}\right) \leq E \log\left(1 + \frac{Y}{\epsilon}\right) \leq E\frac{Y}{\epsilon} < \infty,$$

where Y is the uniform integrable upper bound for our gambles and ϵ is the minimal possible loss lower bound.

By Theorem 1 in Shiryaev (1984), Chapter VII, we conclude that the set $\{S_t \to -\infty\}$ is a null set. Indeed, on the set $\{S_t \to -\infty\}$, S is bounded above. The theorem then states that the limit of S exists and is finite (almost surely), and thus cannot be negative infinity. □

3.4 Time Consistency

An important question arising in a dynamic framework is how the conditional risks at different times are interrelated. This question leads to the important notion of time–consistency.

A dynamic risk measure $(\rho_s)_{s=1\ldots t-1}$ is called time–consistent if for any gamble X_t^1, X_t^2 and for all $s = 1, ..., t-2$ it holds that

$$\rho_{s+1}(X_t^1) \geq \rho_{s+1}(X_t^2) \ a.s. \Longrightarrow \rho_s(X_t^1) \geq \rho_s(X_t^2) \ a.s. \tag{3.3}$$

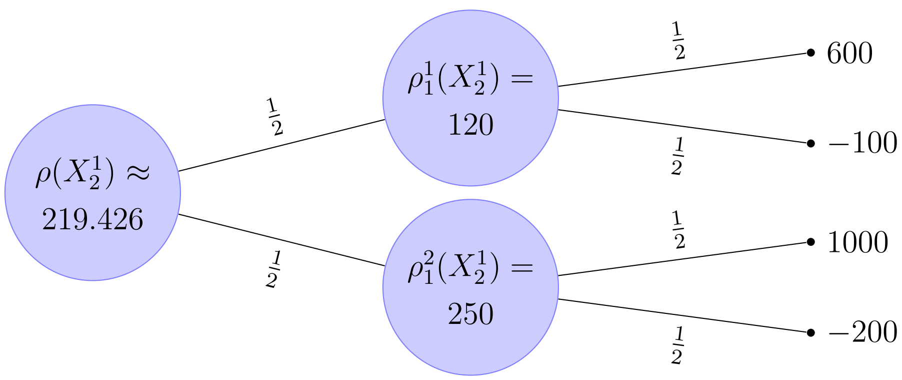
That means, in particular, that if we know that tomorrow in every state of the world gamble X_t^1 is assigned to have a higher risk than gamble X_t^2, this should also hold true today.

Riedel (2004) and Roorda et al. (2005) give representation theorems for time–consistent dynamic coherent risk measures. Similar results for convex risk measures have been obtained by Detlefsen and Scandolo (2005). However, as the Foster–Hart measure of riskiness is neither coherent nor convex we cannot simply conclude from these conditions whether or not the Foster–Hart measure of riskiness is time–consistent.

We therefore create an example that shows that the dynamic Foster–Hart measure of riskiness fails the time–consistency condition.

Example 3 *Consider two discrete gambles X_2^1 and X_2^2 that have their payments in two periods $(t = 2)$ from now. They are distributed according to the binomial trees given below. In $t = 1$, two states of the world are possible which occur with equal probability $\frac{1}{2}$. We compute the riskiness today $(t = 0)$ and in each state in $t = 1$.*

Gamble X_2^1 has the following structure:

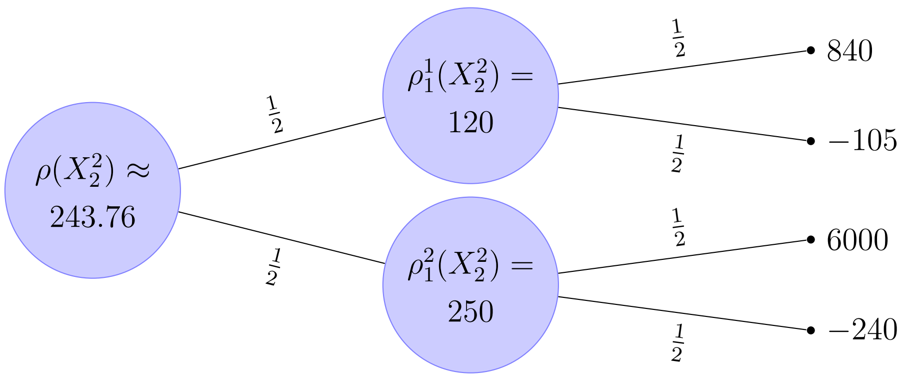


Hence, X_2^1 has the payoffs $\{600, -100, 1000, -200\}$ occurring each with equal probability. The riskiness in $t = 1$ in state one is the unique positive *solution to the equation*

$$1/2\log(1+600/\rho_1^1(X_2^1)) + 1/2\log(1-100/\rho_1^1(X_2^1)) = 0.$$

Solving this gives $\rho_1^1(X_2^1) = 120$*. In a similar way we get that the riskiness number in state two is* $\rho_1^2(X_2^1) = 250$ *and that the riskiness today is* $\rho(X_2^1) \approx 219.426$*.*

The second gamble X_2^2 *is distributed according to the following tree:*



Although the payoffs of X_2^2 *differ from the payoffs of* X_2^1*, the riskiness numbers at time* $t = 1$ *coincide. Today, however, the risk of* X_2^2 *is strictly greater than the risk of* X_2^1 *which contradicts the time–consistency condition* (3.3)*. Therefore, the dynamic extended Foster–Hart riskiness is not time–consistent.*

From this example we can see an interesting distinction between the dynamic Foster–Hart measure of riskiness and the original static one. As $E[\log(1+\frac{X}{\rho})|F_0] = E\log(1+\frac{X}{\rho})$, we can view the riskiness in time zero as the static Foster–Hart riskiness number.

Where the conditional Foster–Hart measure of riskiness in time 1 does not differentiate between gamble X_1 and X_2, the unconditional one does. In fact, the conditional measure at time 1 requires for both gambles in each state the same wealth to overcome the possibility of going bankrupt, whereas the static risk measure requires a higher wealth for the second gamble.

This may lead to an inconsistent behavior. Suppose, for instance, an agent's strategy is solely induced by the Foster–Hart riskiness measure (i.e. the agent rejects a gamble whenever her wealth is below the Foster–Hart number and accepts it otherwise) and her wealth is \$230. At $t = 1$ both gambles would be rejected in the lower node and might be accepted in the upper one. At time 0, however, the agent would reject X^2 but accept X^1. Nevertheless, as Theorem (3) shows, both measures entail the same no–bankruptcy property.

3.5 Existence of the Foster–Hart Index

In this section, we show that our concept of the extended Foster–Hart index of riskiness is well–defined. All inequalities and equalities between random variables are assumed to hold $P-$almost surely.

Let X_t be a gamble for the σ–field $\mathcal{F}_s$. Without loss of generality, we can assume $L_s(X_t) = 1$ almost surely (else replace X_t by $X_t/L_s(X_t)$).

We write $\mathcal{G} = \mathcal{F}_s$ and $X = X_t$ in the following for shorter notation. We fix a regular version $P(\tilde{\omega}, d\omega)$ for the conditional probability distribution of X given $\mathcal{G}$ (which exists as X takes values in a Polish space). Whenever we write conditional expectations or probabilities in the following, we have this regular version in mind.

Firstly, we argue that there exists no solution to equation (3.2) on $\mathcal{B}$. For $\tilde{\omega} \in \mathcal{B}$, we consider the function

$$\lambda \mapsto \int_\Omega \log(1 + \lambda X(\omega)) P(\tilde{\omega}, d\omega).$$

We can now apply the arguments given in chapter 2. Indeed, one can easily verify that this function is concave and positive for all $0 < \lambda \leq 1$. This shows that no solution to the defining equation exists on $\mathcal{B}$.

Let us move on to the proof of Theorem (2). For our construction, we need that there exist wealth levels W for which we accept the gamble X given $\mathcal{G}$. We thus start with the following observation.

Lemma 3 *There exist $\mathcal{G}$–measurable random variables $W \geq 1$ such that*

$$E[\log(W + X)|\mathcal{G}] \geq \log W\,.$$

In particular, this holds true for all W with

$$W \geq \frac{2E[X^2|\mathcal{G}]}{E[X|\mathcal{G}]}$$

and

$$\left|\frac{X}{W}\right| \leq \frac{1}{2}\,.$$

Proof of Lemma 3: We use the estimate

$$\log(1+x) \geq x - 2x^2 \tag{3.4}$$

for $|x| \leq 1/2$ (which one can obtain from a Taylor–expansion and the Lagrange version of the error term). Take an $\mathcal{G}$–measurable W with

$$W \geq \frac{2E[X^2|\mathcal{G}]}{E[X|\mathcal{G}]}$$

and

$$\left|\frac{X}{W}\right| \leq \frac{1}{2}\,.$$

Such W exists because X has finite variance. For example, we can take

$$W = \max\left\{\frac{2E[X^2|\mathcal{G}]}{E[X|\mathcal{G}]}, 2|X|, 1\right\}\,.$$

As $|X/W| \leq 1/2$, $\log(1+X/W)$ is everywhere defined. By the estimate (3.4), we obtain

$$\begin{aligned} E\left[\log\left(1+\frac{X}{W}\right)|\mathcal{G}\right] &\geq E\left[\frac{X}{W} - \frac{2X^2}{W^2}|\mathcal{G}\right] \\ &= \frac{1}{W}\left(E\left[X|\mathcal{G}\right] - \frac{E\left[2X^2|\mathcal{G}\right]}{W}\right)\,, \end{aligned}$$

and now we can use the fact that $W \geq 2E[X^2|\mathcal{G}]/E[X|\mathcal{G}]$ to conclude that

$$E\left[\log\left(1+\frac{X}{W}\right)|\mathcal{G}\right] \geq 0\,.$$

□

As a consequence of the preceding lemma, the set

$$\Lambda = \{\lambda\ \mathcal{G} - \text{measurable} |\, 0 < \lambda \leq 1, E\left[\log(1 + \lambda X)\,|\mathcal{G}\right] \geq 0\}$$

is not empty. Let λ_0 be the $\mathcal{G}$–essential supremum of Λ. By definition, λ_0 is $\mathcal{G}$–measurable and $\lambda_0 \geq \lambda$ for all $\lambda \in \Lambda$. Moreover, λ_0 is the smallest $\mathcal{G}$–measurable random variable with these properties.

The set Λ is upwards directed: take $\lambda_1, \lambda_2 \in \Lambda$. Then we have for $\lambda_3 = \max\{\lambda_1, \lambda_2\}$

$$\begin{aligned} E\left[\log(1 + \lambda_3 X)|\mathcal{G}\right] &= 1_{\{\lambda_1 \geq \lambda_2\}} E\left[\log(1 + \lambda_1 X)|\mathcal{G}\right] \\ &\qquad + 1_{\{\lambda_1 < \lambda_2\}} E\left[\log(1 + \lambda_2 X)|\mathcal{G}\right] \\ &\geq 0\,. \end{aligned}$$

The other properties being obvious, we conclude $\lambda_3 \in \Lambda$. Hence, Λ is upwards directed; as a consequence, there exists a sequence $(\lambda_n) \subset \Lambda$ with $\lambda_n \uparrow \lambda_0$.

Our next claim is $E\left[\log(1 + \lambda_0 X)|\mathcal{G}\right] \geq 0$. The sequence

$$Z_n = -\log\left(1 + \lambda_n X\right)$$

is bounded from below by $-\log(1 + |X|) \geq -|X| \in L^1$. We can then apply Fatou's lemma to conclude

$$-E\left[\log(1 + \lambda_0 X)|\mathcal{G}\right] = E\left[\lim Z_n|\mathcal{G}\right] \leq \liminf -E\left[\log\left(1 + \lambda_n X_n\right)|\mathcal{G}\right] \leq 0\,,$$

or

$$E\left[\log(1 + \lambda_0 X)|\mathcal{G}\right] \geq 0\,.$$

We claim now that we have

$$E\left[\log\left(1 + \lambda_0 X\right)|\mathcal{G}\right] = 0 \tag{3.5}$$

on the set $\{\lambda_0 < 1\}$. This will conclude the proof of our lemma.

It is enough to establish the claim on all sets

$$\Gamma_n = \left\{ \lambda_0 \le 1 - \frac{1}{n} \right\}$$

for all $n \in \mathbb{N}$. From now on, we work on this set only without stating it explicitly. Let

$$A_{m,n} = \left\{ E\left[\log\left(1 + \lambda_0 X\right) | \mathcal{G}\right] \ge \frac{1}{m} \right\} \cap \Gamma_n \,.$$

We will show that $A_{m,n}$ is a null set for all $m, n \in \mathbb{N}$.

Let $\epsilon = 1/(1 + mn)$ and set $\lambda_1 = (1 - \epsilon)\lambda_0 + \epsilon$. Then we have $\lambda_1 > \lambda_0$ and $\lambda_1 \le (1 - \epsilon)(1 - 1/n) + \epsilon = 1 - 1/n + \epsilon/n < 1$. We also note

$$\lambda_1 - \lambda_0 = \epsilon(1 - \lambda_0) \le \epsilon. \tag{3.6}$$

We have

$$1 + \lambda_1 X \ge 1 - \lambda_1 \ge \frac{1 - \epsilon}{n} > 0 \,. \tag{3.7}$$

Thus, $\log(1 + \lambda_1 X)$ is finite (on Γ_n where we work).

We now want to show

$$E\left[\log\left(1 + \lambda_1 X\right) | \mathcal{G}\right] \ge 0$$

on $A_{m,n}$. If $A_{m,n}$ was not a null set, this would contradict the definition of λ_0 as the $\mathcal{G}$–essential supremum of Λ.

In order to establish the desired inequality, it is enough to show

$$E\left[\log\left(1 + \lambda_1 X\right) | \mathcal{G}\right] - E\left[\log\left(1 + \lambda_0 X\right) | \mathcal{G}\right] \ge -\frac{1}{m}$$

because of the definition of $A_{m,n}$. Now, on the set $\{X \ge 0\}$ we have $\log\left(1 + \lambda_1 X\right) \ge \log\left(1 + \lambda_0 X\right)$.

We need a uniform estimate for $\log\left(1 + \lambda_1 X\right) - \log\left(1 + \lambda_0 X\right)$ on the set

$\{X < 0\}$. With the help of the mean value theorem, we obtain on $\{X < 0\}$

$$\log\left(\frac{1+\lambda_1 X}{1+\lambda_0 X}\right) \geq -\frac{n}{1-\epsilon}\,(\lambda_1 - \lambda_0) \geq -\frac{n\epsilon}{1-\epsilon}\,.$$

(By the mean value theorem and (3.6), we have

$$\log(1+\lambda_1 X) - \log(1+\lambda_0 X) = \frac{1}{\xi}(\lambda_1 - \lambda_0)X$$

for some ξ in between $1+\lambda_1 X$ and $1+\lambda_0 X$. By (3.7) $1+\lambda_1 X \geq (1-\epsilon)/n$. Hence, we have $0 < 1/\xi \leq n/(1-\epsilon)$.)

By the definition of ϵ, we thus have

$$\log\left(\frac{1+\lambda_1 X}{1+\lambda_0 X}\right) \geq -\frac{n\epsilon}{1-\epsilon} = -\frac{1}{m}$$

uniformly on $\{X < 0\}$ as desired. It follows

$$\begin{aligned}
&E\left[\log(1+\lambda_1 X)\,|\mathcal{G}\right] - E\left[\log(1+\lambda_0 X)\,|\mathcal{G}\right] \\
&\geq E\left[\log\left(\frac{1+\lambda_1 X}{1+\lambda_0 X}\right) 1_{\{X<0\}}|\mathcal{G}\right] \\
&\geq -\frac{1}{m}\,.
\end{aligned}$$

Chapter 4

Fear of the Market or Fear of the Competitor? Ambiguity in a Real Options Game [1]

4.1 Introduction

Since the seminal contribution of Chen and Epstein (2002), there has been a solid framework for dealing with Gilboa and Schmeidler (1989) max–min preferences in a continuous time multiple prior model of ambiguity. This model has been applied to several problems in economics and finance to gain valuable insights in the consequences of a form of Knightian uncertainty, as opposed to risk, on economic decisions. The main insight of Chen and Epstein (2002) is that in order to find the max–min value of a payoff stream under a particular kind of ambiguity (called strongly rectangular) we need to identify the upper-rim generator of the set of multiple priors, and value the payoff stream as if this is the true process governing the payoffs.

In the literature this process has become known as the *worst–case prior*, because it identifies the prior that at any given time t gives the lowest expected discounted payoff from time t. In the literature on investment under uncertainty (so–called "real options") the approach has been used to value

[1] Parts of this chapter can be found in Hellmann and Thijssen (2015).

investment projects when the decision maker is not sure about the stochastic process governing the cash-flows resulting from the project. Typically, this literature models cash–flows as geometric Brownian motions and ambiguity takes the form of κ-ambiguity over the true trend of the diffusion. In that case it has been shown by Nishimura and Ozaki (2007) that the worst–case at any time t corresponds to the lowest possible trend that is considered under κ-ambiguity.

In this chapter, we extend the Nishimura and Ozaki model to a timing game between two firms, which both have the option to invest in a project, where one firm is ambiguous about the process governing cash-flows, and the other firm (potentially) has a cost disadvantage. For our analysis, however, the assumption that only one firm is ambiguous does not play a role. In fact section 6 shows that our result can easily be adopted to the case where both firms are ambiguous, possibly to a different degree. This assumption is made to illustrate the difference an introduction of ambiguity makes compared to a purely risky world. In our analysis we may now compare a risky firm to an ambiguous one.

In such timing games, players typically have to balance the expected future payoffs of being the first or second firm to invest; the leader and follower roles, respectively.

The purpose of this chapter is threefold. Firstly, we want to explore the effects of ambiguity on the leader and follower payoffs to players. Secondly, we wish to extend the equilibrium concepts for stochastic timing games[2] to include ambiguous players. Thirdly, we want to investigate the interaction of ambiguity and cost (dis-) advantages on equilibrium investment scenarios.

Our main conclusions are as follows. First, contrary to all of the literature on ambiguity in the real options literature, the worst–case prior is not always the lowest possible trend under κ–ambiguity. As in any timing game, an ambiguous player has to consider the payoffs of the leader and follower roles.

[2] Since the seminal contribution of Fudenberg and Tirole (1985) for deterministic timing games, many attempts to defining equilibria in stochastic timing games has been made such as Thijssen (2010), Thijssen et al. (2012), de Villemeur et al. (2014), Boyarchenko and Levendorski (2014), Azevedo and Paxson (2014), Huisman and Kort (2015).

The payoffs of the latter role follow along very similar lines as in Nishimura and Ozaki (2007), i.e. the worst–case payoff corresponds to valuing the follower's payoff stream as if the payoffs are driven by the diffusion with the lowest admissible trend under κ-ambiguity. For the leader's payoff, however, the situation is different, because of the interplay between two opposing forces. On the one hand, the leader's payoff consists of current payoffs of being the leader. The worst–case for these payoffs is represented by the lowest admissible trend, like in the follower payoff. There is, however, another force at work: the risk that the other player invests as well, which reduces the firm's monopoly payoff to a duopoly payoff. This event has a downward effect on the leader's payoff and is discounted using the expected time it takes until the other firm enters the market. This expected time is reached faster for higher values of the trend of the stochastic process, so that the worst–case for this part of the leader's payoffs is represented by the highest admissible trend. We use an analysis based on backward stochastic differential equations and "g–expectations", as introduced by Peng (1997), to study which effect dominates. It turns out that for small values of the stochastic process, the worst–case always corresponds to the lowest admissible trend, whereas for higher values the highest admissible trend may represent the worst–case, depending on the underlying parameters.

Secondly, we show that equilibria can be of two types. First, there may be preemptive equilibria in which one of the firms invests at a time where it is not optimal for either firm to do so. This type of equilibrium is familiar from the literature (e.g. Fudenberg and Tirole (1985), Weeds (2002), Pawlina and Kort (2006)) but we use a technique recently developed by Riedel and Steg (2014) to rigorously prove existence of this type of equilibrium rather than relying on fairly ad hoc arguments that are often used in the existing literature. It should be pointed out here that in a preemptive equilibrium it is known a.s. ex ante which firm is going to invest first. This firm will invest at a point in time where its leader value exceeds its follower value, but where its competitor is indifferent between the two roles. A second type of equilibrium that can exist is a sequential equilibrium, in which one firm invests at a time where it is optimal for them to do so. By that we mean that the firm would choose the same time to invest even if it knew that the other firm could not preempt. Each game always has at least an equilibrium of one of these two types, which can not co-exist. These two types of equilibrium each lead to a clear prediction, a.s., as to which firm invests first. The role of first mover depends crucially on the levels of ambiguity and cost (dis-) advantage, as we show in a numerical analysis.

As mentioned above we obtain our equilibrium results by using techniques developed by Riedel and Steg (2014). It should be pointed out that we cannot simply adopt their strategies to our setting due to the presence of an ambiguous player. In fact, the notion of extended mixed strategy as introduced in Riedel and Steg (2014) presents a conceptual problem here. An extended mixed strategy consists, in essence, of a distribution over stopping times as well as a coordination device that allows players to coordinate in cases where equilibrium considerations require one and only one firm to invest and it is not clear a priori which firm this should be. In our model we need this coordination device as well, but we do not want ambiguity to extend to the uncertainty created by this coordination mechanism, i.e. ambiguity is over payoffs exclusively. This presents problems if we want to define payoffs to the ambiguous firm if it plays a mixture over stopping times. For equilibrium existence, however, such mixtures are not needed, so we choose to restrict attention to what we call *extended pure strategies*, which consist of a stopping time and an element related to the coordination mechanism mentioned above. By making this simplifying assumption, together with strong rectangularity of the set of priors, we can write the worst–case payoff of a pair of extended pure strategies as a sum of worst–cases of leader and follower payoffs.

At the end of this chapter, we provide some comparative statics. We explore the effect a change of the degree of ambiguity, the volatility and the cost–disadvantage has on equilibrium outcomes. We show numerically that the investment thresholds of the ambiguous firm increase with the degree of uncertainty. Due to the construction of the set of priors via κ–ignorance, an increase of volatility implies also an increase of uncertainty. Both firm's investment thresholds rise with the volatility. The effect on the set of priors, however, yields that the thresholds of the ambiguous firm is more affected by a change of the volatility. Finally, it is shown that ambiguity might outbalance the cost–disadvantage. Pawlina and Kort (2006) argued that in a purely risky world, the low–cost firm always becomes the leader. We show that this might change if ambiguity is introduced.

4.2 The Model

We follow Pawlina and Kort (2006) in considering two firms that are competing to implement a new technology. Uncertainty in the market is modeled on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ using a geometric Brownian motion

$$\frac{dX}{X} = \mu dt + \sigma dB,$$

where $(B_t)_{t\geq 0}$ is a Wiener process. The sunk costs of investment are $I > 0$ for firm 1 and αI, $\alpha \geq 1$, for firm 2. So, firm 1 may have a cost advantage.

The payoff streams are given by processes $(D_{k\ell} X_t)_{t\geq 0}$, where $D_{k\ell}$, $k, \ell = 0, 1$, denotes the scaling factor if the firm's investment status is k ($k = 0$ if the firm has not invested and $k = 1$ if the firm has invested) and the investment status of the competitor is ℓ. It is assumed that $D_{10} > D_{11} \geq D_{00} \geq D_{01} \geq 0$, and that there is a first mover advantage, i.e. $D_{10} - D_{00} > D_{11} - D_{01}$.

We assume that firm 1 may have a cost advantage, but also that it is ambiguous about the trend μ. Following the recent literature on ambiguity in continuous time models, we model the set of priors that the firm considers using a set of density generators. The set of measures that is considered by the firm is denoted by $\mathscr{P}^{\Theta}$, where Θ is a set of density generators. A process $(\theta_t)_{t\geq 0}$ is a density generator if the process $\left(M_t^{\theta}\right)_{t\geq 0}$, where

$$\frac{dM_t^{\theta}}{M_t^{\theta}} = -\theta_t dB_t, \quad M_0^{\theta} = 1, \tag{4.1}$$

is a P–martingale. Such a process $(\theta_t)_{t\geq 0}$ generates a new measure P^{θ} via the Radon–Nikodym derivative $d\mathsf{P}^{\theta}/d\mathsf{P} = M_{\infty}^{\theta}$.

In order to use density generators as a model for ambiguity the set Θ needs some more structure. Following Chen and Epstein (2002), the set of density generators, Θ, is chosen as follows. Let $(\Theta_t)_{t\geq 0}$ be a collection of correspondences $\Theta_t : \Omega \twoheadrightarrow \mathbb{R}$, such that

1. There is a compact subset $K \subset \mathbb{R}$, such that $\Theta_t(\omega) \subseteq K$, for all $\omega \in \Omega$ and all $t \in [0, T]$;
2. For all $t \in [0, T]$, Θ_t is compact-valued and convex-valued;
3. For all $t \in (0, T]$, the mapping $(s, \omega) \mapsto \Theta_s(\omega)$, restricted to $[0, t] \times \Omega$, is $\mathcal{B}[0, t] \times \mathcal{F}_t$-measurable;
4. $0 \in \Theta_t(\omega)$, $dt \otimes d\mathsf{P}$-a.e.

The set of density generators is then taken to be,

$$\Theta = \{(\theta_t)_{t\geq 0} \, | \theta_t(\omega) \in \Theta_t(\omega), d\mathsf{P} - \text{a.e., all } t \geq 0\},$$

and the resulting set of measures $\mathscr{P}^\Theta$ is called *strongly-rectangular*. For sets of strongly rectangular priors the following has been obtained by Chen and Epstein (2002):

1. $\mathsf{P} \in \mathscr{P}^\Theta$;
2. All measures in $\mathscr{P}^\Theta$ are uniformly absolutely continuous with respect to P and are equivalent to P;
3. For every $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathsf{P})$, there exists $\mathsf{P}^* \in \mathscr{P}^\Theta$ such that for all $t \geq 0$,

$$\mathsf{E}^{\mathsf{P}^*}[X|\mathcal{F}_t] = \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q[X|\mathcal{F}_t]. \tag{4.2}$$

Finally, for further reference, define the *upper-rim generator* $(\theta_t^*)_{t\geq 0}$, where

$$\theta_t^* = \arg\max\{\sigma_w(t)\theta_t | \theta_t \in \Theta_t\}. \tag{4.3}$$

Note that $(\theta_t^*)_{t\geq 0} \in \Theta$.

From Girsanov's theorem it immediately follows that under $\mathsf{P}^\theta \in \mathscr{P}^\Theta$, the process $\left(B_t^\theta\right)_{t\geq 0}$, defined by

$$B_t^\theta = B_t + \int_0^t \theta_s ds,$$

is a P^θ-Brownian motion and that, under P^θ, the process $(X_t)_{t\geq 0}$ follows the diffusion

$$\frac{dX_t}{X_t} = \mu^\theta(t)dt + \sigma dB_t^\theta.$$

Furthermore,

$$\mu^\theta(t) = \mu - \sigma\theta_t.$$

In the remainder we will assume that $\Theta_t = [-\kappa, \kappa]$, for all $t > 0$, for some $\kappa > 0$. Denote $\Delta = [\underline{\mu}, \overline{\mu}] = [\mu - \sigma\kappa, \mu + \sigma\kappa]$. This form of ambiguity is called *κ–ignorance* (cf. Chen and Epstein (2002)). The advantages of using this definition of ambiguity are that (i) Θ is strongly rectangular so that the results stated above apply and (ii) the upper–rim generator takes a convenient form, namely $\theta_t^* = \kappa$, for all $t \geq 0$. In addition, it can easily be shown that $\left(B_t^\theta\right)_{t\geq 0}$ is a P-martingale for every $(\theta_t)_{t\geq 0} \in \Theta$.

Notice, Cheng and Riedel (2013) show that $\kappa-$ignorance can be applied in an infinite time–horizon. In particular, they show that value functions taken under drift ambiguity in the infinite time horizon are nothing but the limits of value functions of finite time horizons T if $T \to \infty$.

In our model, we assume firm 1 to be ambiguity averse in the sense of Gilboa and Schmeidler (1989).

For our upcoming computations it is crucial to assume that any finite threshold will be hit by the underlying stochastic process with probability one given any possible drift $\mu \in [\underline{\mu}, \overline{\mu}]$. For a geometric Brownian motion this is ensured if $\underline{\mu} \geq 1/2\sigma^2$.

Finally, the discount rate is assumed to be $r > \overline{\mu}$.

4.3 Leader and Follower Value Functions

4.3.1 The Non–Ambiguous Firm

Assume firm 1 becomes the leader at t. Then the non–ambiguous firm 2 solves the optimal stopping problem

$$F_2(x_t) = \sup_{\tau_2^F \geq t} \mathsf{E}^{\mathsf{P}} \left[\int_t^{\tau_2^F} e^{-r(s-t)} D_{01} X_s ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s - e^{-r(\tau_2^F - t)} \alpha I \Big| \mathcal{F}_t \right]. \tag{4.4}$$

Thus, τ_2^F is the optimal time firm 2 invests as a follower.

On the other hand, if the non–ambiguous firm becomes the leader at a certain point in time t, its value function is

$$L_2(x_t) = \mathsf{E}^{\mathsf{P}} \left[\int_t^{\tau_1^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_1^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - \alpha I \Big| \mathcal{F}_t \right], \tag{4.5}$$

where τ_1^F denotes the optimal time at which the ambiguous firm invests as a follower. From the standard literature on real option games (cf. Pawlina and Kort (2006)) we know that the former value function can be written as

$$F_2(x_t) = \begin{cases} \frac{x_t D_{01}}{r-\mu} + \left(\frac{x_2^F (D_{11} - D_{01})}{r-\mu} - \alpha I \right) \left(\frac{x_t}{x_2^F} \right)^{\beta(\mu)}, & \text{if} \quad x_t \leq x_2^F, \\ \frac{x_t D_{11}}{r-\mu} - \alpha I & \text{if} \quad x_t > x_2^F, \end{cases} \tag{4.6}$$

where τ_2^F is the first hitting time of x_2^F, i.e

$$\tau_2^F = \inf\{s \geq t | X_s \geq x_2^F\}.$$

The standard procedure of dynamic programming yields that the threshold x_2^F is given by

$$x_2^F = \frac{\beta(\mu)}{\beta(\mu) - 1} \frac{\alpha I (r-\mu)}{D_{11} - D_{01}},$$

where $\beta(\mu)$ is the positive root of the fundamental quadratic $1/2\sigma^2\beta(\mu)(\beta(\mu)-$

$1) + \mu\beta(\mu) - r = 0$, which is

$$\beta(\mu) = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}} > 1.$$

Applying the standard techniques of backward induction and dynamic programming, one can show that the leader value (4.5) turns out to be

$$L_2(x_t) = \begin{cases} \frac{x_t D_{10}}{r-\mu} - \alpha I + \frac{x_1^F(D_{11}-D_{10})}{r-\mu}\left(\frac{x_t}{x_1^F}\right)^{\beta(\mu)}, & \text{if } \; x_t \leq x_1^F, \\ \frac{x_t D_{11}}{r-\mu} - I, & \text{if } \; x_t > x_1^F. \end{cases}$$

Accordingly, the real value x_1^F describes the optimal time for the ambiguous firm to become the follower, i.e.

$$\tau_1^F = \inf\{s \geq t | X_s \geq x_1^F\}.$$

What remains to study is the case when both firms invest simultaneously at t. One can show that the shared value function of firm 2 turns out to be

$$M_2(x_t) = \mathsf{E}^{\mathsf{P}}\left[\int_t^\infty e^{-r(s-t)} D_{11} X_s ds - \alpha I \Big| \mathcal{F}_t\right] = \frac{x_t D_{11}}{r-\mu} - \alpha I.$$

4.3.2 The Ambiguous Firm

If ambiguity is introduced, the standard techniques for computing the value functions are not applicable any longer. In our case, where ambiguity is modeled by a strongly rectangular set of density generators, one needs, in contrast to the standard case, to allow for changing priors over time.

The value functions of the ambiguous firm 1 are given by

$$F_1(x) := \sup_{\tau_1^F \geq t} \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q\left[\int_t^{\tau_1^F} e^{-r(s-t)} D_{01} X_s ds + \int_{\tau_1^F}^\infty e^{-r(s-t)} D_{11} X_s - e^{-r(\tau_1^F - t)} I \Big| \mathcal{F}_t\right] \tag{4.7}$$

and

$$L_1(x_t) = \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \Big| \mathcal{F}_t \right] - I, \tag{4.8}$$

respectively.

If the set of priors $\mathscr{P}^\Theta$ is strongly rectangular, it turns out that problem (4.7) can be reduced to a standard optimal stopping problem and, hence, can be solved by using standard techniques. This reduction is possible due to the following lemma, the proof of which is standard and is, thus, omitted.

Lemma 4 *Let $\mathscr{P}^\Theta$ be strongly–rectangular. Then*

$$F_1(x_t) = \sup_{\tau_1^F \geq t} E^{P^{\theta^*}} \left[\int_t^{\tau_1^F} e^{-r(s-t)} D_{01} X_s ds + \int_{\tau_1^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - e^{-r(\tau_1^F - t)} I \Big| \mathcal{F}_t \right], \tag{4.9}$$

where $(\theta_t^*)_{t \geq 0}$ is the upper–rim generator (4.3).

Hence, for the follower problem of the ambiguous firm, the worst–case is always induced by the worst possible drift $\underline{\mu}$. This observation indeed makes sense; the actions of the opponent have, essentially, no influence of the decision as a follower. The problem therefore reduces to one of a "monopolistic" decision maker. Nishimura and Ozaki (2007) already showed that for such decisions, the worst–case is always given by the worst possible trend $\underline{\mu}$.

In other words, we find that the follower value of the ambiguous firm can be expressed by

$$F_1(x_t) = \begin{cases} \frac{x_t D_{01}}{r - \underline{\mu}} + \left(\frac{x_1^F (D_{11} - D_{01})}{r - \underline{\mu}} - I \right) \left(\frac{x_t}{x_1^F} \right)^{\beta(\underline{\mu})}, & \text{if } \; x_t \leq x_1^F, \\ \frac{x_t D_{11}}{r - \underline{\mu}} - I & \text{if } \; x_t > x_1^F, \end{cases} \tag{4.10}$$

where

$$x_1^F = \frac{\beta(\underline{\mu})}{\beta(\underline{\mu}) - 1} \frac{I(r - \underline{\mu})}{D_{11} - D_{01}}.$$

In the similar way, one can argue that for simultaneous investment the value function of firm 1 is induced by the worst–case $\underline{\mu}$ and therefore

$$M_1(x_t) = \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^\infty e^{-r(s-t)} D_{11} X_s ds - I \Big| \mathcal{F}_t \right] = \frac{x_t D_{11}}{r - \underline{\mu}} - I.$$

Determining the leader value function of the ambiguous firm, however, is a different issue. The action of the opponent (in this case the decision when to invest as a follower) is crucial for the computation of the leader function which might lead, as we will see, to a non–trivial behavior of the worst–case prior.

The next theorem describes the leader value function of the ambiguous firm. Two cases are distinguished there. If the difference $D_{10} - D_{11}$ is sufficiently small, we find that the worst–case is, as before, always induced by $\underline{\mu}$. In case this condition is not satisfied, the worst–case is given by $\underline{\mu}$ for values x_t up to a certain threshold x^*, where it jumps to $\overline{\mu}$. The intuition for this fact can already be derived from equation (4.8); the lowest trend $\underline{\mu}$ gives the minimal values for the payoff stream $(D_{kl}X_t)$. However, the higher the trend μ the sooner the stopping time τ_2^F is expected to be reached. The higher payoff stream $(D_{10}X_t)$ is then sooner replaced by the lower one $(D_{11}X_t)$. If the drop of the payoffs becomes sufficiently small, the former effect always dominates the latter. In this case the worst–case is given by $\underline{\mu}$ for each x_t.

Theorem 4 *The worst–case for the leader function of the ambiguous firm is always given by the worst possible drift $\underline{\mu}$ if and only if the following condition holds*

$$\frac{D_{10} - D_{11}}{D_{10}} \leq \frac{1}{\beta_1(\underline{\mu})}. \tag{4.11}$$

In this case, the leader function becomes

$$L_1(x_t) = \begin{cases} \frac{D_{10}x_t}{r-\underline{\mu}} - \left(\frac{x_t}{x_2^F}\right)^{\beta_1(\underline{\mu})} \frac{D_{11}-D_{10}}{r-\underline{\mu}} x_2^F - I & \text{if } x_t < x_2^F \\ \frac{D_{11}x_t}{r-\underline{\mu}} - I & \text{if } x_t \geq x_2^F. \end{cases} \tag{4.12}$$

On the other hand, if $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu})}$, then there exists a unique threshold x^* such that $\underline{\mu}$ is the worst–case on the set $\{X_t < x^*\}$ and $\overline{\mu}$ is the worst–case on $\{x^* \leq X_t < x_2^F\}$. Furthermore, in this case the leader value function is given *by*

$$L_1(x_t) = \begin{cases} \frac{D_{10}x_t}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{x_t}{x^*}\right)^{\beta_1(\underline{\mu})} - I & \textit{if } x_t < x^* \\ \frac{D_{10}x_t}{r-\overline{\mu}} + \frac{(x^*)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} & \\ \cdot\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F & \\ + \frac{(x_2^F)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} & \\ \cdot\left(\left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x^* - I & \textit{if } x^* \leq x_t < x_2^F \\ \frac{D_{11}x_t}{r-\underline{\mu}} - I & \textit{if } x_t \geq x_2^F, \end{cases} \tag{4.13}$$

where $\beta_1(\mu)$ *and* $\beta_2(\mu)$ *are the positive, respective negative root of the quadratic equation* $1/2\sigma^2\beta(\mu)(\beta(\mu)-1) + \mu\beta(\mu) - r = 0$.

In case the worst–case is not trivially given by the lowest possible trend, the value function seems to become a bit messy. However, the terms

$$\frac{(x^*)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \text{ and } \frac{(x_2^F)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}$$

admit a clear interpretation; they represent the expected discount factor that x_2^F is reached before x^* and vice versa, respectively.

Figure 4.1 depicts the implications of Theorem (4). In case the drop of the payoff from being the only one who has invested to the situation that both players have invested is sufficiently big, the value x^* nicely distinguishes between two different "regimes".

For $x_t < x^*$ the ambiguous player fears most a worse development of the underlying stochastic (e.g. the development of the market), that is of the underlying trend μ, whereas at times where $x_t > x^*$ the fear of not being the only one having invested dominates the former one. Indicating this observation, we call the two regimes "fear of the market" and "fear of competition", respectively.

For the proof of Theorem (4), we need a completely different approach compared to the standard literature on real option games. We use back-

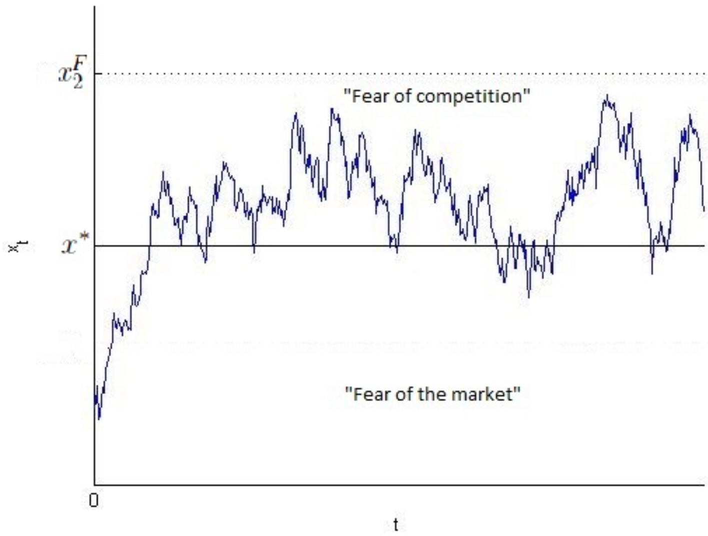


Figure 4.1: The critical value x^* differentiates between two “regimes”.

ward stochastic differential equations[3] and $g-$expectations introduced by Peng (1997). The advantage of this approach lies in the fact that we know the value of our problem at the entry point of the follower. This value yields the starting point for a backward stochastic differential equation. The non–linear Feynman–Kac formula reduces the problem to solving a certain non–linear partial differential equation. From this PDE we are eventually able to derive the worst–case prior.

PROOF: Denote

$$Y_t := \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \middle| \mathcal{F}_t \right].$$

Applying the time consistency property of a rectangular set of density generators gives

$$Y_t = \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \middle| \mathcal{F}_t \right]$$

[3]A brief introduction into the theory of backward stochastic differential equations is given in the Appendix.

$$
\begin{aligned}
&= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q\Bigg[\inf_{Q'\in\mathscr{P}^\Theta}\mathsf{E}^{Q'}\Bigg[\int_t^{\tau_2^F} e^{-r(s-t)}D_{10}X_s ds \\
&\qquad\qquad + \int_{\tau_2^F}^\infty e^{-r(s-t)}D_{11}X_s ds\Big|\mathcal{F}_{\tau_2^F}\Bigg]\Big|\mathcal{F}_t\Bigg] \\
&= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q\Bigg[\int_t^{\tau_2^F} e^{-r(s-t)}D_{10}X_s ds \\
&\qquad + e^{-r(\tau_2^F-t)}\inf_{Q'\in\mathscr{P}^\Theta}\mathsf{E}^{Q'}\left[\int_{\tau_2^F}^\infty e^{-r(s-\tau_2^F)}D_{11}X_s ds\Big|\mathcal{F}_{\tau_2^F}\right]\Big|\mathcal{F}_t\Bigg] \\
&= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q\left[\int_t^{\tau_2^F} e^{-r(s-t)}D_{10}X_s ds + e^{-r(\tau_2^F-t)}\Phi(x_{\tau_2^F})\Big|\mathcal{F}_t\right],
\end{aligned}
$$

where

$$
\Phi(x_t) := \inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\left[\int_t^\infty e^{-r(s-t)}D_{11}X_s ds\Big|\mathcal{F}_t\right] = \frac{D_{11}x_t}{r-\underline{\mu}}. \tag{4.14}
$$

Chen and Epstein (2002) show that Y_t solves the BSDE

$$
-dY_t = g(Z_t)dt - Z_t dB_t,
$$

for the *generator*

$$
g(z) = -\kappa|z| - rY_t + X_t D_{10}.
$$

The boundary condition is given by

$$
Y_{\tau_2^F} = \Phi(x_2^F).
$$

Notice that the stopping time τ_2^F is hit at some finite time $t<\infty$ with probability one by the underlying process (due to our assumption that $\underline{\mu} \geq 1/2\sigma^2$).

In the terminology of Peng (1997), we say that the leader value is the $g-$expectation of the random variable $\Phi(x_2^F)$, and denote it by

$$
Y_t = \mathsf{E}_g[\Phi(x_2^F)|\mathcal{F}_t].
$$

Denote the present value of the leader payoff by L, i.e.

$$L(x_t) = Y_t.$$

The non–linear Feynman–Kac formula[4] (cf. Appendix, Theorem (8)) implies that L solves the non–linear PDE

$$\mathcal{L}_X L(x) + g(\sigma x L'(x)) = 0.$$

Hence, L solves

$$\frac{1}{2}\sigma^2 x^2 L''(x) + \mu x L'(x) - \kappa\sigma x\,|L'(x)| - rL(x) + D_{10}x = 0. \tag{4.15}$$

Expression (4.15) implies that $\underline{\mu}$ is the worst–case on the set $\{x \leq x_2^F | L'(x) > 0\}$ and $\overline{\mu}$ is the worst–case on $\{x \leq x_2^F | L'(x) < 0\}$.

The unique viscosity solution to the PDE (4.15) is given by

$$L(\mu, x) = \frac{D_{10}x}{r-\mu} + Ax^{\beta_1(\mu)} + Bx^{\beta_2(\mu)}, \tag{4.16}$$

where μ equals either $\underline{\mu}$ or $\overline{\mu}$. The constants A and B are determined by some boundary conditions.

One can easily see that for x close to zero we have $L'(x) > 0$. Now two cases are possible; either $L'(x) > 0$ for all $x \in [0, x_2^F]$ or we can find (at least) one point x^* at which the worst–case changes from $\underline{\mu}$ to $\overline{\mu}$.

Let us first assume that $\underline{\mu}$ is always the worst–case. Since $\beta_2(\mu) < 0$, we have $B = 0$. In order to determine the constant A, we apply a value matching condition at x_2^F that gives

$$L(\underline{\mu}, x_2^F) = \frac{D_{10}x_2^F}{r-\underline{\mu}} + A_1 {x_2^F}^{\beta_1(\underline{\mu})} = \frac{D_{11}x_2^F}{r-\underline{\mu}}.$$

[4] Note that Peng (1991) shows that the non–linear Feynman–Kac formula not only holds for deterministic times but also first exit times like τ_2^F.

This implies

$$A_1 = \frac{D_{10} - D_{11}}{r - \underline{\mu}} {x_2^F}^{1-\beta_1(\underline{\mu})},$$

and therefore

$$L(x_t) = \frac{D_{10} x_t}{r - \underline{\mu}} + \left(\frac{x_t}{x_2^F}\right)^{\beta_1(\underline{\mu})} \frac{D_{11} - D_{10}}{r - \underline{\mu}} x_2^F. \tag{4.17}$$

We get that $\underline{\mu}$ is always the worst–case on $[0, x_2^F]$ if and only if $L'(x) \geq 0$ for all $x \leq x_2^F$ (see equation (4.15)). Due to the continuity and concavity of the value function (4.17), this is equivalent to the condition

$$L'(x_2^F) \geq 0.$$

Therefore,

$$\begin{aligned} L'(x_2^F) = \frac{D_{10}}{r - \underline{\mu}} + \left(\frac{D_{11} - D_{10}}{r - \underline{\mu}}\right) \beta_1(\underline{\mu}) \left(\frac{x_2^F}{x_2^F}\right)^{\beta_1(\underline{\mu})-1} &\geq 0 \\ D_{11} - D_{10} &\geq -\frac{D_{10}}{\beta_1(\underline{\mu})} \\ \frac{D_{10} - D_{11}}{D_{10}} &\leq \frac{1}{\beta_1(\underline{\mu})}. \end{aligned}$$

If this condition is not satisfied, the worst–case changes at some point $x^* < x_2^F$ from $\underline{\mu}$ to $\overline{\mu}$, where x^* is determined by the condition $L'(x^*) = 0$. We denote by $\tilde{L_1}(\underline{\mu}, x)$ the solution to (4.16) on $[0, x^*]$ and by $\hat{L_1}(\overline{\mu}, x)$ the solution to (4.16) on $[x^*, x_2^F]$. The unknowns in equation (4.16) are determined by applying twice a value matching condition and once a smooth pasting condition. Indeed, it must hold that

1. $\hat{L_1}(\overline{\mu}, x_2^F) = \Phi(x_2^F)$,
2. $\tilde{L_1}(\underline{\mu}, x^*) = \hat{L_1}(\overline{\mu}, x^*)$,
3. $\tilde{L_1}'(\underline{\mu}, x^*) = \hat{L_1}'(\overline{\mu}, x^*)$.

In case $\underline{\mu}$ is not always the worst–case, the unique solution of (4.16) is given by

$$L(x_t) = 1_{x_t<x^*}\tilde{L_1}(\overline{\mu}, x_t) + 1_{x_t\geq x^*}\hat{L_1}(\overline{\mu}, x_t),$$

where[5]

$$\tilde{L_1}(\underline{\mu}, x_t) = \frac{D_{10}x_t}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{x_t}{x^*}\right)^{\beta_1(\underline{\mu})},$$

and

$$\hat{L_1}(\overline{\mu}, x_t) =$$

$$\frac{D_{10}x_t}{r-\overline{\mu}} + \frac{(x^*)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F$$
$$+ \frac{(x_2^F)^{\beta_1(\overline{\mu})}x_t^{\beta_2(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})}x_t^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x^*.$$

We can easily verify that $\hat{L_1}$ and $\tilde{L_1}$ satisfy the boundary conditions. Indeed,

$$\hat{L_1}(\overline{\mu}, x_2^F) =$$

$$\frac{D_{10}x_2^F}{r-\overline{\mu}} + \frac{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F$$
$$+ \frac{(x_2^F)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x^*$$
$$=\frac{D_{10}x_2^F}{r-\overline{\mu}} + \left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F$$
$$=\frac{D_{11}x_2^F}{r-\overline{\mu}}$$
$$=\Phi(x_2^F),$$

[5]In the Appendix a detailed derivation of the functions $\tilde{L_1}$ and $\hat{L_1}$ is presented.

and

$$
\begin{aligned}
&\hat{L_1}(\overline{\mu}, x^*) = \\
&\quad \frac{D_{10}x^*}{r-\overline{\mu}} + \frac{(x^*)^{\beta_2(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x^*)^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x_2^F \\
&\quad + \frac{(x_2^F)^{\beta_1(\overline{\mu})}(x^*)^{\beta_2(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \left(\left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x^* \\
&= \frac{D_{10}x^*}{r-\overline{\mu}} + \left(\left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x^* \\
&= \frac{D_{10}x^*}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}} \\
&= \tilde{L_1}(\underline{\mu}, x^*).
\end{aligned}
$$

To prove the smooth pasting condition at x^* requires a bit more work. Firstly, we observe that the value x^* is chosen such that it always holds that $\tilde{L_1'}(\underline{\mu}, x^*) = 0$.

The next lemma shows that there exists such a value x^*, which is unique and satisfies also $\hat{L_1}(\overline{\mu}, x^*) = 0$.

Lemma 5 *If $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu})}$, then there exists one and only one value x^* that solves $\hat{L_1'}(\overline{\mu}, x^*) = 0$ on $(0, x_2^F]$.*

The proof is reported in the Appendix.

□

Remark 2 *The leader value function L_1 is always concave on $[0, x_2^F]$ even if the worst–case changes at some point. We prove this fact in the Appendix.*

Figure 4.2 shows a typical run of the leader and follower value functions of both the ambiguous and the non–ambiguous firm. We observe that the leader value function of firm 1 drops below its follower value function if x_t is close to x_2^F. The reason for that is that x_1^F and x_2^F differ (in the illustrated case we have $x_2^F < x_1^F$). That means that the leader and follower value functions

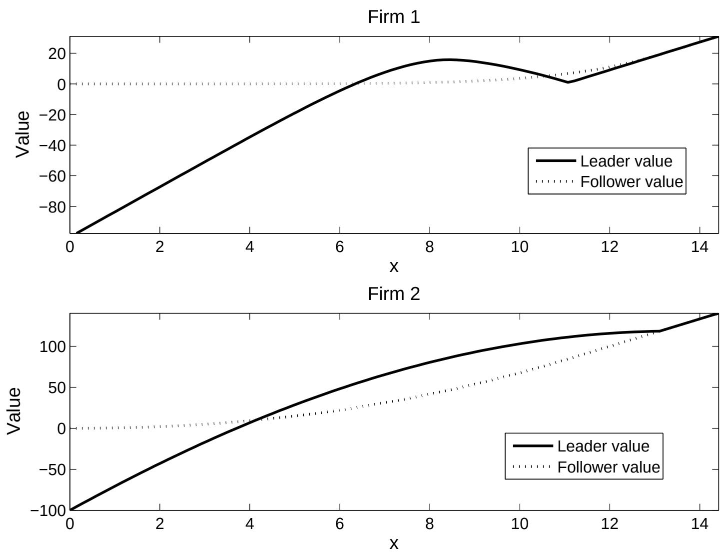


Figure 4.2: The leader and follower value functions of the ambiguous and non–ambiguous firm.

hit the shared value function M at different times. This is the case because x_1^F and x_2^F are determined using a different trend. But even if firms use the same prior, in some cases we would observe this patter, namely if we consider cost–asymmetric firms, i.e. $\alpha > 1$.

4.3.3 Optimal Leader Threshold

Next we want to determine the optimal time to invest as a leader. Suppose firm 2 knows it becomes the leader and searches for the optimal time to invest. It then faces at time t the following optimal stopping problem

$$L^*(x_t) = \sup_{\tau_{L,2}^t \geq t} \mathsf{E}^{\mathsf{P}} \Bigg[\int_t^{\tau_{L,2}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,2}^t}^{\tau_1^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_1^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - e^{-r(\tau_{L,2}^t - t)} \alpha I \Big| \mathcal{F}_t \Bigg].$$

The solution can be found by applying the standard techniques and is well known from the literature. It is given by

$$\tau_{L,2}^{t} = \inf\{s \geq t | X_s \geq x_2^L\},$$

where

$$x_2^L = \frac{\beta_1(\mu)}{\beta_1(\mu) - 1} \frac{\alpha I(r - \mu)}{D_{10} - D_{00}}.$$

The ambiguous firm solves the following optimal stopping problem

$$L^*(x_t) = \sup_{\tau_{L,1}^t \geq t} \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds \right.$$
$$\left. + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds - e^{-r(\tau_{L,1}^t - t)} I \Big| \mathcal{F}_t \right].$$

Again, in order to determine this stopping time for the ambiguous firm, we cannot apply the standard procedure. Nevertheless, the stopping time does not differ from the one of a non–ambiguous firm given a drift $\underline{\mu}$.

Proposition 2 *The optimal time to invest as a leader for the ambiguous firm is*

$$\tau_{L,1}^{t} = \inf\{s \geq t | X_s \geq x_1^L\},$$

where

$$x_1^L = \frac{\beta_1(\underline{\mu})}{\beta_1(\underline{\mu}) - 1} \frac{I(r - \underline{\mu})}{D_{10} - D_{00}}.$$

For the proof we refer to the Appendix.

4.4 Equilibrium Analysis

The appropriate equilibrium concept for a game with ambiguity as described here is not immediately clear. In this chapter, we consider two types of equilibria: *preemptive equilibria* in which firms try to preempt each other at some times where it is sub–optimal to invest, and *sequential equilibria*, where one firm invests at its optimal time.

4.4.1 Strategies and Payoffs

The appropriate notion of subgame perfect equilibrium for our game is developed in Riedel and Steg (2014). Let $\mathscr{T}$ denote the set of stopping times with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. The set $\mathscr{T}$ will act as the set of (pure) strategies. Given the definitions of the leader, follower and shared payoffs above, the timing game is

$$\Gamma = \left\langle (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}), \mathscr{P}^\Theta, \mathscr{T} \times \mathscr{T}, (L_i, F_i, M_i)_{i=1,2}, (\pi_i)_{i=1,2} \right\rangle,$$

where, for $(\tau_1, \tau_2) \in \mathscr{T} \times \mathscr{T}$,

$$\begin{aligned}\pi_1 &= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q[L_1 1_{\tau_1<\tau_2} + F_1 1_{\tau_1>\tau_2} + M_1 1_{\tau_1=\tau_2}], \quad \text{and} \\ \pi_2 &= \mathsf{E}^\mathsf{P}[L_2 1_{\tau_1>\tau_2} + F_2 1_{\tau_1<\tau_2} + M_2 1_{\tau_1=\tau_2}].\end{aligned}$$

The subgame starting at stopping time $\vartheta \in \mathscr{T}$ is the tuple

$$\Gamma^\vartheta = \left\langle (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P}), \mathscr{P}^\Theta, \mathscr{T}_\vartheta \times \mathscr{T}_\vartheta, (L_i, F_i, M_i)_{i=1,2}, (\pi_i^\vartheta)_{i=1,2} \right\rangle,$$

where $\mathscr{T}_\vartheta$ is the set of stopping times no smaller than ϑ a.s.,

$$\mathscr{T}_\vartheta := \{\tau \in \mathscr{T} | \tau \geq \vartheta, \mathsf{P} - a.s.\},$$

and, for $(\tau_1, \tau_2) \in \mathscr{T}_\vartheta \times \mathscr{T}_\vartheta$,

$$\begin{aligned}\pi_1^\vartheta &= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q[L_1 1_{\tau_1<\tau_2} + F_1 1_{\tau_1>\tau_2} + M_1 1_{\tau_1=\tau_2} | \mathcal{F}_\vartheta], \quad \text{and} \\ \pi_2^\vartheta &= \mathsf{E}^\mathsf{P}[L_2 1_{\tau_1>\tau_2} + F_2 1_{\tau_1<\tau_2} + M_2 1_{\tau_1=\tau_2} | \mathcal{F}_\vartheta].\end{aligned}$$

As it is argued in Riedel and Steg (2014), careful consideration has to be given to the appropriate notion of strategy. They show that the notion of extended mixed strategy is versatile and intuitively appealing. For the subgame Γ^ϑ this is a pair of processes $(G^\vartheta, \alpha^\vartheta)$, both taking values in $[0, 1]$, with the following properties.

1. G^ϑ is adapted, has right–continuous and non–decreasing sample paths,

with $G^\vartheta(s) = 0$ for all $s < \vartheta$, $\mathsf{P} - a.s.$

2. α^ϑ is progressively measurable with right–continuous sample paths whenever its value is in $(0,1)$, $\mathsf{P} - a.s.$

3. On $\{t \geq \vartheta\}$, it holds that

$$\alpha^\vartheta(t) > 0 \Rightarrow G^\vartheta(t) = 1, \quad \mathsf{P} - a.s.$$

We use the convention that

$$G^\vartheta(0-) \equiv 0, \quad G^\vartheta(\infty) \equiv 1, \quad \text{and} \quad \alpha^\vartheta(\infty) \equiv 1.$$

For our purposes extended mixed strategies are, in fact, more general than necessary. Therefore, we will restrict attention to what we will call *extended pure strategies*. For the subgame Γ^ϑ this is a pair of extended mixed strategies $(G_i^\vartheta, \alpha_i^\vartheta)_{i=1,2}$, where G_i^ϑ is restricted to take values in $\{0,1\}$. In other words, in an extended pure strategy a firm does not mix over stopping times, but potentially mixes over its "investment intensity" α^ϑ.

An extended pure strategy for the game Γ is then a collection $(G^\vartheta, \alpha^\vartheta)_{\vartheta\in\mathscr{T}}$ of extended pure strategies in subgames Γ^ϑ, $\vartheta \in \mathscr{T}$ satisfying the time consistency conditions that for all $\vartheta, \nu \in \mathscr{T}$ it holds that

1. $\nu \leq t \in \mathbb{R}_+ \Rightarrow G^\vartheta(t) = G^\vartheta(\nu-) + (1 - G^\vartheta(\nu-))G^\nu(t)$, P-a.s. on $\{\vartheta \leq \nu\}$,

2. $\alpha^\vartheta(\tau) = \alpha^\nu(\tau)$, P-a.s., for all $\tau \in \mathscr{T}$.

The importance of the α component in the definition of extended pure strategy becomes obvious in the definition of payoffs. Essentially α allows both for immediate investment and coordination between firms. It leads to investment probabilities that can be thought of as the limits of conditional stage investment probabilities of discrete–time behavioral strategies with vanishing period length. In the remainder, let $\hat{\tau}_i^\vartheta$ be the first time that α_i^ϑ is strictly positive, and let $\hat{\tau}^\vartheta$ be the first time that at least one α^ϑ is non–zero in the subgame Γ^ϑ, i.e.

$$\hat\tau_i^\vartheta = \inf\{t \geq \vartheta | \alpha_i^\vartheta(t) > 0\}, \quad \text{and} \quad \hat\tau^\vartheta = \inf\{t \geq \vartheta | \alpha_1^\vartheta(t) + \alpha_2^\vartheta(t) > 0\},$$

respectively. At time $\hat\tau^\vartheta$ the extended pure strategies induce a probability measure on the state space

$$\Lambda = \{\{\text{Firm 1 becomes the leader}\}, \{\text{Firm 2 becomes the leader}\},$$
$$\{\text{Both firms invest simultaneously}\}\},$$

for which we will use the shorthand notation

$$\Lambda = \{(L,1), (L,2), M\}.$$

Riedel and Steg (2014) show that the probability measure on Λ, induced by the pair $(\alpha_1^\vartheta, \alpha_1^\vartheta)$, is given by

$$\lambda_{L,i}^\vartheta(\hat\tau^\vartheta) = \begin{cases} \frac{\alpha^\vartheta_{i,\hat\tau^\vartheta}(1-\alpha^\vartheta_{j,\hat\tau^\vartheta})}{\alpha^\vartheta_{i,\hat\tau^\vartheta}+\alpha^\vartheta_{j,\hat\tau^\vartheta}-\alpha^\vartheta_{i,\hat\tau^\vartheta}\alpha^\vartheta_{j,\hat\tau^\vartheta}} & \text{if } \hat\tau_i^\vartheta = \hat\tau_j^\vartheta \\ & \text{and } \alpha_i^\vartheta(\hat\tau_i^\vartheta), \alpha_j^\vartheta(\hat\tau_i^\vartheta) > 0 \\ 1 & \text{if } \hat\tau_i^\vartheta < \hat\tau_j^\vartheta, \text{ or } \hat\tau_i^\vartheta = \hat\tau_j^\vartheta \\ & \text{and } \alpha_j^\vartheta(\hat\tau_j^\vartheta) = 0 \\ 0 & \text{if } \hat\tau_i^\vartheta > \hat\tau_j^\vartheta, \text{ or } \hat\tau_i^\vartheta = \hat\tau_j^\vartheta \\ & \text{and } \alpha_j^\vartheta(\hat\tau_j^\vartheta) = 0 \\ \frac{1}{2}\Big(\liminf_{t\downarrow\hat\tau_i^\vartheta} \frac{\alpha_i^\vartheta(t)(1-\alpha_j^\vartheta(t))}{\alpha_i^\vartheta(t)+\alpha_j^\vartheta(t)-\alpha_i^\vartheta(t)\alpha_j^\vartheta(t)} & \text{if } \hat\tau_i^\vartheta = \hat\tau_j^\vartheta, \\ +\limsup_{t\downarrow\hat\tau_i^\vartheta} \frac{\alpha_i^\vartheta(t)(1-\alpha_j^\vartheta(t))}{\alpha_i^\vartheta(t)+\alpha_j^\vartheta(t)-\alpha_i^\vartheta(t)\alpha_j^\vartheta(t)}\Big) & \alpha_i^\vartheta(\hat\tau_i^\vartheta) = \alpha_j^\vartheta(\hat\tau_j^\vartheta) = 0, \\ & \text{and } \alpha_i^\vartheta(\hat\tau_i^\vartheta+), \alpha_j^\vartheta(\hat\tau_j^\vartheta+) > 0, \end{cases}$$

and

$$\lambda_M^\vartheta(\hat{\tau}^\vartheta) = \begin{cases} 0 & \text{if } \hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta,\ \alpha_i^\vartheta(\hat{\tau}_i^\vartheta) = \alpha_j^\vartheta(\hat{\tau}_i^\vartheta) = 0, \\ & \text{and } \alpha_i^\vartheta(\hat{\tau}_i^\vartheta+), \alpha_j^\vartheta(\hat{\tau}_i^\vartheta+) > 0 \\ \frac{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta \alpha_{j,\hat{\tau}^\vartheta}^\vartheta}{\alpha_{i,\hat{\tau}^\vartheta}^\vartheta + \alpha_{j,\hat{\tau}^\vartheta}^\vartheta - \alpha_{i,\hat{\tau}^\vartheta}^\vartheta \alpha_{j,\hat{\tau}^\vartheta}^\vartheta} & \text{otherwise.} \end{cases}$$

Note the following:

1. If $\hat{\tau}_i^\vartheta < \hat{\tau}_j^\vartheta$ there is no coordination problem: firm i becomes the leader λ-a.s.;
2. If $\hat{\tau}_i^\vartheta = \hat{\tau}_j^\vartheta$, but $\alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0$, there is no coordination problem: firm i becomes the leader λ-a.s.;
3. In the degenerate case where $\alpha_i^\vartheta(\hat{\tau}_i^\vartheta) = \alpha_j^\vartheta(\hat{\tau}_j^\vartheta) = 0$, and $\alpha_i^\vartheta(\hat{\tau}_i^\vartheta+), \alpha_j^\vartheta(\hat{\tau}_j^\vartheta+) > 0$, the leader role is effectively assigned on the basis of the flip of a fair coin;
4. There is no ambiguity (for firm 1) over the measure λ.

In order to derive the payoffs to firms, let $\tau_{G,i}^\vartheta$ denote the first time that G_i^ϑ jumps to one, i.e.

$$\tau_{G,i}^\vartheta = \inf\{t \geq \vartheta | G_i^\vartheta(t) > 0\}.$$

The payoff to the ambiguous firm of a pair of extended pure strategies $((G_1, \alpha_1), (G_2, \alpha_2))$ in the subgame Γ^ϑ is given by

$$\begin{aligned} &V_1^\vartheta(G_1^\vartheta, \alpha_1^\vartheta, G_2^\vartheta, \alpha_2^\vartheta) := \\ &\inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \Bigg[1_{\tau_{G,1}^\vartheta < \min\{\tau_{G,2}^\vartheta, \hat{\tau}^\vartheta\}} \Bigg(\int_\vartheta^{\tau_{G,1}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds + \int_{\tau_{G,1}^\vartheta}^{\tau_2^F} e^{-r(s-\vartheta)} D_{10} X_s ds \\ &\qquad + \int_{\tau_2^F}^\infty e^{-r(s-\vartheta)} D_{11} X_s ds - e^{-r(\tau_{G,1}^\vartheta - \vartheta)} I \Bigg) \Bigg| \mathcal{F}_\vartheta \Bigg] \\ &+ \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \Bigg[1_{\tau_{G,2}^\vartheta < \min\{\tau_{G,1}^\vartheta, \hat{\tau}^\vartheta\}} \Bigg(\int_\vartheta^{\tau_{G,2}^\vartheta} e^{-r(s-\vartheta)} D_{00} X_s ds + \int_{\tau_{G,2}^\vartheta}^{\tau_1^F} e^{-r(s-\vartheta)} D_{01} X_s ds \end{aligned}$$

$$+\int_{\tau_1^F}^{\infty} e^{-r(s-\vartheta)}D_{11}X_s - e^{-r(\tau_1^F-\vartheta)}I\Big)\Big|\mathcal{F}_\vartheta\Big]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\Big[1_{\tau_{G,1}^\vartheta=\tau_{G,2}^\vartheta<\hat\tau^\vartheta}\Big(\int_\vartheta^{\tau_{G,1}^\vartheta} e^{-r(s-\vartheta)}D_{00}X_s ds$$

$$+\int_{\tau_{G,1}^\vartheta}^{\infty} e^{-r(s-\vartheta)}D_{11}X_s ds\Big)\Big|\mathcal{F}_\vartheta\Big]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\Big[1_{\hat\tau^\vartheta\le\min\{\tau_{G,1}^\vartheta,\tau_{G,1}^\vartheta\}}\lambda_{L,1}^\vartheta(\hat\tau^\vartheta)\Big(\int_\vartheta^{\hat\tau^\vartheta} e^{-r(s-\vartheta)}D_{00}X_s ds$$

$$+\int_{\hat\tau^\vartheta}^{\tau_2^F} e^{-r(s-\vartheta)}D_{10}X_s ds+\int_{\tau_2^F}^{\infty} e^{-r(s-\vartheta)}D_{11}X_s ds - e^{-r(\tau_{G,1}^\vartheta-\vartheta)}I\Big)\Big|\mathcal{F}_\vartheta\Big]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\Big[1_{\hat\tau^\vartheta\le\min\{\tau_{G,1}^\vartheta,\tau_{G,1}^\vartheta\}}\lambda_{L,2}^\vartheta(\hat\tau^\vartheta)\Big(\int_\vartheta^{\hat\tau^\vartheta} e^{-r(s-\vartheta)}D_{00}X_s ds$$

$$+\int_{\hat\tau^\vartheta}^{\tau_1^F} e^{-r(s-\vartheta)}D_{01}X_s ds+\int_{\tau_1^F}^{\infty} e^{-r(s-\vartheta)}D_{11}X_s - e^{-r(\tau_1^F-\vartheta)}I\Big)\Big|\mathcal{F}_\vartheta\Big]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\Big[1_{\hat\tau^\vartheta\le\min\{\tau_{G,1}^\vartheta,\tau_{G,1}^\vartheta\}}\lambda_{M}^\vartheta(\hat\tau^\vartheta)\Big(\int_\vartheta^{\hat\tau^\vartheta} e^{-r(s-\vartheta)}D_{00}X_s ds$$

$$+\int_{\hat\tau^\vartheta}^{\infty} e^{-r(s-\vartheta)}D_{11}X_s ds\Big)\Big|\mathcal{F}_\vartheta\Big].$$

Hence, the payoff of the ambiguous firm can written as

$$V_1^\vartheta(G_1^\vartheta,\alpha_1^\vartheta,G_2^\vartheta,\alpha_2^\vartheta) := \inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\left[1_{\tau_{G,1}^\vartheta<\min\{\tau_{G,2}^\vartheta,\hat\tau^\vartheta\}}L_1(x_\vartheta)\Big|\mathcal{F}_\vartheta\right]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\left[1_{\tau_{G,2}^\vartheta<\min\{\tau_{G,1}^\vartheta,\hat\tau^\vartheta\}}F_1(x_\vartheta)\Big|\mathcal{F}_\vartheta\right]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\left[1_{\tau_{G,1}^\vartheta=\tau_{G,2}^\vartheta<\hat\tau^\vartheta}M_1(x_\vartheta)\Big|\mathcal{F}_\vartheta\right]$$

$$+\inf_{Q\in\mathscr{P}^\Theta}\mathsf{E}^Q\left[1_{\hat\tau^\vartheta\le\min\{\tau_{G,1}^\vartheta,\tau_{G,1}^\vartheta\}}\lambda_{L,1}^\vartheta(\hat\tau^\vartheta)L_1(x_\vartheta)\Big|\mathcal{F}_\vartheta\right]$$

$$+ \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q \left[1_{\hat{\tau}^\vartheta \leq \min\{\tau^\vartheta_{G,1},\tau^\vartheta_{G,1}\}} \lambda^\vartheta_{L,2}(\hat{\tau}^\vartheta) F_1(x_\vartheta) \middle| \mathcal{F}_\vartheta \right]$$

$$+ \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q \left[1_{\hat{\tau}^\vartheta \leq \min\{\tau^\vartheta_{G,1},\tau^\vartheta_{G,1}\}} \lambda^\vartheta_{M}(\hat{\tau}^\vartheta) M_1(x_\vartheta) \middle| \mathcal{F}_\vartheta \right].$$

In the same way, the payoff for the unambiguous firm can be written as

$$\begin{aligned}
V_2^\vartheta(G_2^\vartheta,\alpha_2^\vartheta,G_1^\vartheta,\alpha_1^\vartheta) :=& \mathsf{E}^\mathsf{P}\left[1_{\tau^\vartheta_{G,2}<\min\{\tau^\vartheta_{G,2},\hat{\tau}^\vartheta\}} L_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right] \\
&+\mathsf{E}^\mathsf{P}\left[1_{\tau^\vartheta_{G,1}<\min\{\tau^\vartheta_{G,2},\hat{\tau}^\vartheta\}} F_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right] \\
&+\mathsf{E}^\mathsf{P}\left[1_{\tau^\vartheta_{G,1}=\tau^\vartheta_{G,2}<\hat{\tau}^\vartheta} M_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right] \\
&+\mathsf{E}^\mathsf{P}\left[1_{\hat{\tau}^\vartheta \leq \min\{\tau^\vartheta_{G,1},\tau^\vartheta_{G,1}\}} \lambda^\vartheta_{L,2}(\hat{\tau}^\vartheta) L_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right] \\
&+\mathsf{E}^\mathsf{P}\left[1_{\hat{\tau}^\vartheta \leq \min\{\tau^\vartheta_{G,1},\tau^\vartheta_{G,1}\}} \lambda^\vartheta_{L,1}(\hat{\tau}^\vartheta) F_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right] \\
&+\mathsf{E}^\mathsf{P}\left[1_{\hat{\tau}^\vartheta \leq \min\{\tau^\vartheta_{G,1},\tau^\vartheta_{G,1}\}} \lambda^\vartheta_{M}(\hat{\tau}^\vartheta) M_2(x_\vartheta)\middle|\mathcal{F}_\vartheta\right].
\end{aligned}$$

4.4.2 Preemptive and Sequential Equilibria

An equilibrium for the subgame Γ^ϑ is a pair of extended pure strategies $\big((\bar{G}_1^\vartheta,\bar{\alpha}_1^\vartheta),(\bar{G}_2^\vartheta,\bar{\alpha}_2^\vartheta)\big)$, such that for each firm $i=1,2$ and every extended pure strategy $(G_i^\vartheta,\alpha_i^\vartheta)$ it holds that

$$V_i^\vartheta(\bar{G}_i^\vartheta,\bar{\alpha}_i^\vartheta,\bar{G}_j^\vartheta,\bar{\alpha}_j^\vartheta) \geq V_i^\vartheta(G_i^\vartheta,\alpha_i^\vartheta,\bar{G}_j^\vartheta,\bar{\alpha}_j^\vartheta),$$

for $j\neq i$. A subgame perfect equilibrium is a pair of extended pure strategies $\big((\bar{G}_1,\bar{\alpha}_1),(\bar{G}_2,\bar{\alpha}_2)\big)$, such that for each $\vartheta\in\mathscr{T}$ the pair $\big((\bar{G}_1^\vartheta,\bar{\alpha}_1^\vartheta),(\bar{G}_2^\vartheta,\bar{\alpha}_2^\vartheta)\big)$ is an equilibrium in the subgame Γ^ϑ.

There are several types of equilibria of interest in this model. Fix $\vartheta\in\mathscr{T}$. For firm i we denote the optimal time of investment, assuming that the other firm cannot preempt, in the subgame Γ^ϑ by $\tau^\vartheta_{L,i}$, i.e.

$$\tau^\vartheta_{L,i} = \inf\{t\geq\vartheta | X_t \geq x_i^L\}.$$

We also define the *preemption region* as the part of the state space where both firms prefer to be the leader rather than the follower, i.e.

$$\mathcal{P} = \{x \in \mathbb{R}_+ | (L_1(x) - F_1(x)) \wedge (L_2(x) - F_2(x)) > 0\}.$$

The first hitting time of $\mathcal{P}$ in the subgame Γ^{ϑ} is denoted by τ_P^{ϑ}.

We distinguish between two different equilibrium concepts. Lemma (6) determines a preemptive equilibrium.

Lemma 6 *(Riedel and Steg (2014)) Suppose $\vartheta \in \mathscr{T}$ satisfies $\vartheta = \tau_P^{\vartheta}$ $\mathsf{P} - a.s.$ Then $\big((G_1^{\vartheta}, \alpha_1^{\vartheta}), (G_2^{\vartheta}, \alpha_2^{\vartheta})\big)$ given by*

$$\alpha_i^{\vartheta}(t) = \begin{cases} 1 & \text{if } t = \tau_P^t,\ L_t^j = F_t^j, \text{ and } (L_t^i > F_t^i \text{ or } F_t^j = M_t^j) \\ 1_{L_t^1 > F_t^1} 1_{L_t^2 > F_t^2} \frac{L_t^j - F_t^j}{L_t^j - M_t^j} & \text{otherwise,} \end{cases}$$

for any $t \in [\vartheta, \infty)$ and $G_i^{\vartheta} = 1_{t \geq \vartheta}$, $i = 1, 2$, $j \in \{1, 2\}$ i, are an equilibrium in the subgame at ϑ.

In a preemptive equilibrium both firms try to preempt each other. Investment takes place sooner than it optimally would, i.e. the time one firm would invest without the fear of being preempted. The resulting equilibrium in the latter case is called sequential equilibrium. For certain underlying parameters, the preemption time τ_P^{ϑ} is greater than the optimal investment time $\tau_{L,i}^{\vartheta}$ of firm i. A sequential equilibrium is given by the next lemma.

Lemma 7 *Suppose $\vartheta = \tau_{L,i}^{\vartheta} < \tau_P^{\vartheta}$ $\mathsf{P} - a.s.$ for one $i \in \{1, 2\}$. Then $\big((G_1^{\vartheta}, \alpha_1^{\vartheta}), (G_2^{\vartheta}, \alpha_2^{\vartheta})\big)$ given by*

$$\alpha_i^{\vartheta}(\vartheta) = 1, G_i^{\vartheta}(t) = 0 \text{ for all } t < \vartheta,\ \ G_j^{\vartheta}(t) = 0 \text{ for all } t \leq \vartheta$$

are an equilibrium in the subgame at ϑ.

Proof: The stopping time $\tau_{L,i}^{\vartheta}$ is determined in Proposition (2) as the stopping time that maximizes the leader payoff. Hence, without the threat of being preempted by its opponent, i.e. $\tau_{L,i}^{\vartheta} < \tau_P^{\vartheta}$ $\mathsf{P} - a.s.$, it is not optimal to deviate for firm i. Firm j does not want to stop before $\tau_{L,i}^{\vartheta}$ as its payoff of becoming the leader is strictly smaller than becoming the follower up to τ_P^{ϑ}. $\square$

Now, we are finally able to formulate a subgame perfect equilibrium for our game.

Theorem 5 *There exists a subgame perfect equilibrium* $((G_1, \alpha_1), (G_2, \alpha_2))$ *with* α_i^{ϑ} *and* G_1^{ϑ} *given by*

(i) Lemma (6) *if either* $\vartheta \geq \tau_P^{\vartheta}$ $\mathsf{P}-$*a.s. or* $\tau_P^{\vartheta} \leq \tau_{L,i}^{\vartheta}$ $\mathsf{P}-$*a.s..*

(ii) Lemma (7) *otherwise (i.e.* $\vartheta < \tau_P^{\vartheta}$ $\mathsf{P}-$*a.s. and* $\tau_P^{\vartheta} > \tau_{L,i}^{\vartheta}$ $\mathsf{P}-$*a.s.).*

Proof: Optimality for case (ii) follows along the same lines as in the proof of Lemma (7).

If $\vartheta \geq \tau_P^{\vartheta}$ $\mathsf{P}-a.s.$, then optimality for case (i) follows directly from Lemma (6). What remains to prove is that, in case $\vartheta < \tau_P^{\vartheta}$ $\mathsf{P}-a.s.$, neither of the firms wants to invest sooner than τ_P^{ϑ}.

We start with firm 2. Suppose that firm 1 plays the preemption equilibrium strategy. Then if firm 2 plays the preemption strategy, its payoff is $V_2(x) = \mathsf{E}_x[e^{-r\tau_P} L_2(x_P)]$, for any $x < x_P$. (This is the case, because, either the other firm is indifferent between the leader and follower role at x_P, in which case firm 2 becomes the leader, or firm 2 is indifferent in which case $F_2(x_P) = L_2(x_P)$.) Note that we have $V_2(x) = \mathsf{E}_x[e^{-r\tau_P} L_2(x_P)] = \left(\frac{x}{x_P}\right)^{\beta_1(\mu)} L_2(x_P)$ (cf. Dixit and Pindyck (1994), Chapter 9, Appendix A). V_2 is a strictly increasing function, with $V_2(x_P) = L_2(x_P)$ and $V_2(0) = 0 > L_2(0)$, so that $V_2(x) > L_2(x)$ for any $x < x_P$.

The only deviations $\hat{\tau}$ that could potentially give a higher payoff have $\hat{\tau} < \tau_P$, P-a.s. Consider the first hitting time $\hat{\tau}$ of some $\hat{x} < x_P$. Let $\hat{V}_2$ denote the payoff to firm 2 of this strategy (while the other firm plays its preemption strategy). For $\hat{x} \leq x < x_P$, it holds that $\hat{V}_2(x) = L_2(x) < V_2(x)$.

For $x < \hat{x}$, note that $\hat{V}_2(x) = \left(\frac{x}{\hat{x}}\right)^{\beta_1(\mu)} L_2(\hat{x}) = \frac{L_2(\hat{x})}{\hat{x}^{\beta_1(\mu)}} x^{\beta_1(\mu)}$. Consider the mapping $x \mapsto \frac{L_2(x)}{x^{\beta_1(\mu)}}$. This function attains its maximum at $x_2^L > x_P$. Therefore, its derivative is positive on $(0, x_P)$, implying that $V_2(x) > \hat{V}_2(x)$. Any

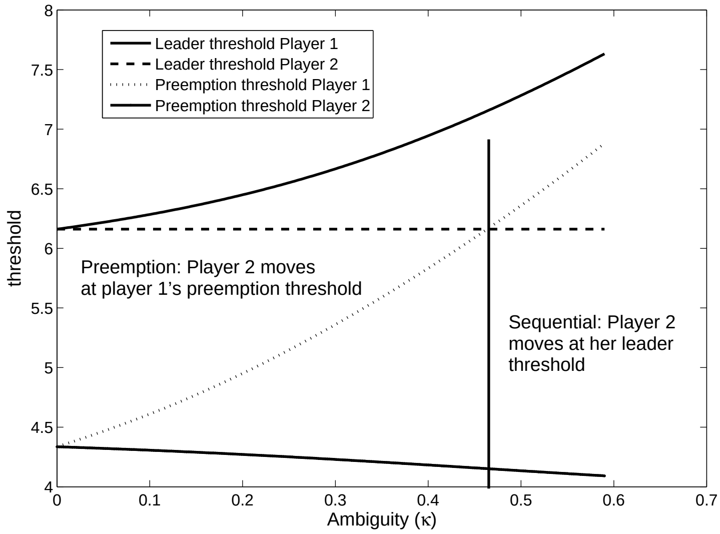


Figure 4.3: The resulting thresholds with respect to κ for the values $D_{10} = 1.8$, $D_{11} = 1$, $D_{01} = 0$, $D_{00} = 0$, $r = 0.1$, $\sigma = 0.1$, $\mu = 0.04$, $I = 100$ and $\alpha = 1$.

stopping time τ can be written as a mixture of first hitting times. So, no stopping time $\hat{\tau}$ with $\hat{\tau} < \tau_P$, P-a.s. yields a higher payoff than τ_P.

For firm 1, the argument is similar after realizing that $V_1(x) = \frac{L_1(x_P)}{x_P^{\beta_1(\underline{\mu})}} x^{\beta_1(\underline{\mu})}$ and $\hat{V}_1(x) = \frac{L_1(\hat{x})}{\hat{x}^{\beta_1(\underline{\mu})}} x^{\beta_1(\underline{\mu})}$. This holds because $x_P < x_1^L < x^*$, so that $\underline{\mu}$ is the trend under the worst–case measure for every $x \in (0, x_P]$.

$\square$

4.5 Comparative Statics

In this section, we analyze the sensitivity of equilibrium outcome with respect to a change of the degree of ambiguity κ, the volatility σ and the cost difference α, respectively.

4.5.1 Comparative Statics With Respect to κ

Nishimura and Ozaki (2007) argued that in a monopolistic model where the firm faces drift uncertainty, an increase in κ postpones investment and decreases the profit.

In our duopoly framework, we observe that both the leader and the follower value function of the ambiguous firm decrease with an increase of κ.

For equilibrium outcomes it is important to investigate how investment times (or thresholds) vary with a change of κ. We find that the follower investment threshold of the ambiguous firm rises if κ increases. Hence, the non–ambiguous firm's payoff increases as it enjoys the benefits of being the only one who has invested for a longer time. Further, we easily see that x_1^L increases with κ.

To see what happens to the preemption time $\tau_P^1 := \inf\{t \geq 0 | L_1(x_t) \geq F_1(x_t)\}$, we need to consider $L_1 - F_1$. Both functions L_1 and F_1 decline by a decrease of κ. However, due to the complexity of the ambiguous firm's leader value function, it is not possible to come up with an analytic result about which function decreases more. For this reason, we consider some numerical examples which suggest that the leader function is more affected by a change of κ than the follower function.

Figure 4.3 depicts the change of the leader thresholds as well as the preemption thresholds of both firms with respect to κ. Starting with complete symmetric firms ($\alpha = 1$ and $\kappa = 0$), Figure 4.3 shows that both the preemption threshold and the leader threshold of firm 1 increases with κ. This indicates that L_1 decreases more by an increase in κ than F_1. This observation makes sense; if it were the other way around, firm 1 could benefit from an increase of kappa. Indeed, if firm 1's preemption threshold would decrease more than firm 2's, firm 1 might benefit by receiving the leader role for ever bigger κ.

4.5.2 Comparative Statics With Respect to σ

Comparative statics with respect to the volatility σ is even more complex as a change of sigma affects not only the volatility but also the interval of possible trends. Since $[\underline{\mu}, \overline{\mu}] = [\mu - \sigma\kappa, \mu + \sigma\kappa]$, an increase in σ enlarges the uncertainty. Notice, a change of σ and a change of κ of the same size have exactly the same impact on the interval of possible trends.

From the standard literature on real options it is well known that an in-

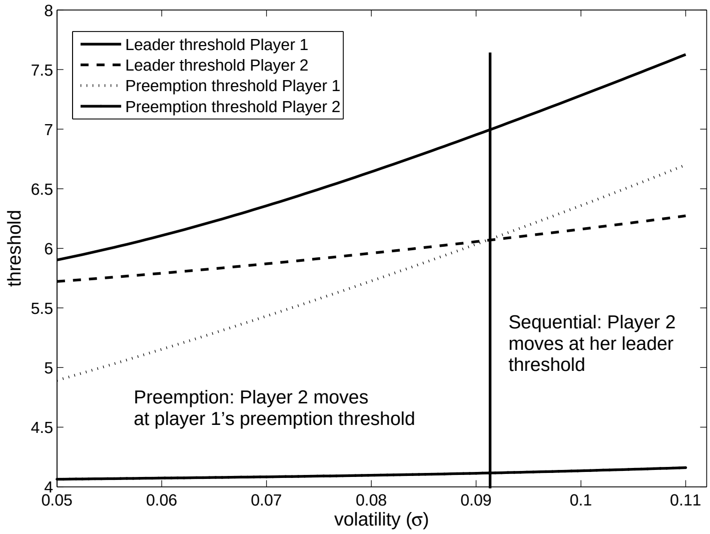


Figure 4.4: The resulting thresholds with respect to σ for the values $D_{10} = 1.8$, $D_{11} = 1$, $D_{01} = 0$, $D_{00} = 0$, $r = 0.1$, $\kappa = 0.5$, $\mu = 0.04$. $I = 100$ and $\alpha = 1$.

crease of σ postpones the investment of a monopolistic firm in a purely risky environment (cf. Nishimura and Ozaki (2007)).

Figure 4.4 shows what happens to the investment thresholds in our framework. All thresholds for both firms increase with the volatility. Due to the effect on the interval of possible trends, however, firm 1's thresholds rise much stronger.

4.5.3 Comparative Statics With Respect to α

Let us move on to investigate the effect the degree of cost asymmetry α has on equilibrium scenarios. In a purely risky framework, the firm that has the lower investment cost always becomes the leader (cf. Pawlina and Kort (2006)). This result, however, might change if ambiguity is introduced. Figure 4.5 shows that even if the the non–ambiguous firm has a higher cost of investment, it might become the leader anyway. Ambiguity, therefore, might outbalance the cost advantage.

From Figure 4.5 we can observe that the preemption threshold as well

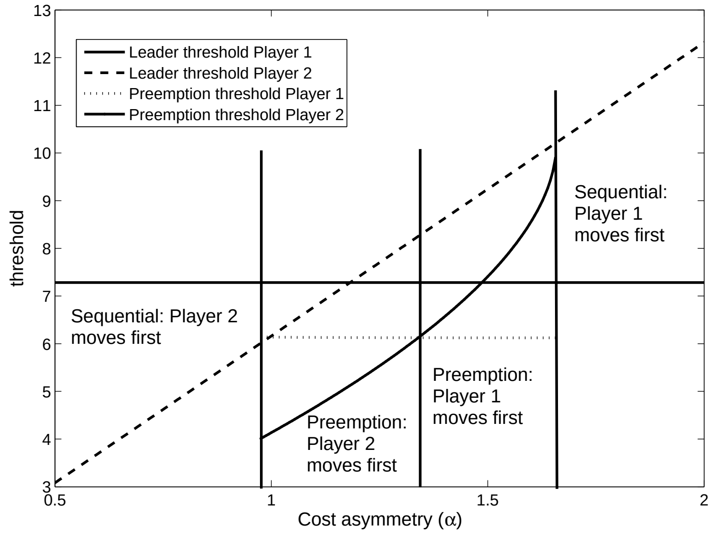


Figure 4.5: The resulting thresholds with respect to α for the values $D_{10} = 1.8$, $D_{11} = 1$, $D_{01} = 0$, $D_{00} = 0$, $r = 0.1$, $\sigma = 0.1$, $\mu = 0.04$, $I = 100$ and $\kappa = 0.5$.

as the leader threshold of firm 2 increase with α. To the far right, there even does not exist a preemption threshold anymore, as the cost disadvantage is too big such that firm 2's leader function always lies below its follower function on $[0, x_1^F]$. Firm 1's leader threshold is unaffected by a change of α. Its preemption threshold, however, is slightly decreasing. The reason for this fact might not be obvious in case condition (4.11) is not satisfied. First note, firm 1's follower function is not affected by a change of α. Further note, the preemption point can only lie in the region where L_1 is increasing. That means, if the worst–case changes at some point, the preemption point is smaller than x^*. Thus, the function needed to be considered is

$$\frac{D_{10}x_t}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{x_t}{x^*}\right)^{\beta_1(\underline{\mu})} - I.$$

This function is also not directly affected by a change of α. Yet, due to the fact that x_2^F increases with α, L_1 increases in the region $[x^*, x_2^F]$. Since the smooth pasting condition has to be fulfilled, this implies that x^* moves to the left. This, however, means that L_1 is also increasing in the region before x^* is

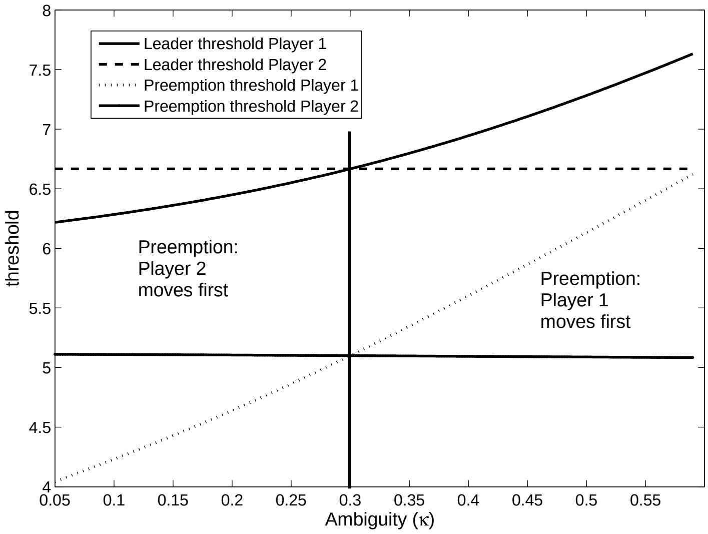


Figure 4.6: The resulting thresholds with respect to κ_2 for the values $D_{10} = 1.8$, $D_{11} = 1$, $D_{01} = 0$, $D_{00} = 0$, $r = 0.1$, $\sigma = 0.1$, $\mu = 0.04$, $\alpha = 1$, $I = 100$ and $\kappa_1 = 0.3$.

reached. This implies that the preemption threshold of firm 1 is decreasing.

4.6 Both Player Ambiguous

We want to emphasize that our analysis is independent of the assumption that only one of the firms is ambiguous. Throughout the paper, this assumption is made in order to elaborate the difference that an introduction of ambiguity makes in contrast to purely risky world.

We may very well allow also for both firms to be ambiguous about the trend of the underlying dynamics. We even do not need to require that the firms have the same degree of ambiguity (same κ).

In fact, for the analysis of the worst–case prior, it is only required that the degree of ambiguity and the cost of investment of each player is common knowledge (such that each firm is able to compute the follower threshold of her competitor). The determination of the follower and leader value function of a second ambiguous firm would be completely analogues to that we presented in section 3.2. Recall that ambiguity is assumed not to be about strategies but about payoffs exclusively. This implies, knowing the new value functions, the equilibrium analysis follows along the same lines as presented in section 4.

In Figure 4.6, we draw different equilibrium outcomes for the case that both players are ambiguous, possibly with a different degree. The firms are assumed to be symmetric in terms of the investment costs. The degree of ambiguity for firm 1 is $\kappa_1 = 0.3$. We vary the degree of ambiguity for the second firm. We see that both the preemption threshold and the leader threshold of firm 2 are strictly increasing, whereas the preemption threshold as well as the leader threshold of firm 1 are slightly decreasing.

4.7 Appendix

4.7.1 Derivation of $\tilde{L}_1$ and $\hat{L}_1$

In the proof of our main theorem, we only verified that the given functions $\tilde{L}_1$ and $\hat{L}_1$ are indeed a solution to the differential equation (4.16). Here, we demonstrate how one can derive from the given value matching and smooth pasting conditions the resulting functions $\tilde{L}_1$ and $\hat{L}_1$.

PROOF: Firstly, we apply the value matching condition at x_2^F, i.e.

$$\frac{D_{10} x_2^F}{r - \overline{\mu}} + A_1 (x_2^F)^{\beta_1(\overline{\mu})} + B (x_2^F)^{\beta_2(\overline{\mu})} = \frac{D_{11} x_2^F}{r - \underline{\mu}}.$$

Hence,

$$A_1 = \left(\frac{D_{11}}{r - \mu} - \frac{D_{10}}{r - \overline{\mu}} - B (x_2^F)^{\beta_2(\overline{\mu}) - 1} \right) (x_2^F)^{1 - \beta_1(\overline{\mu})}. \tag{4.18}$$

Secondly, we apply the value matching condition at x^*, i.e.

$$\frac{D_{10} x^*}{r - \underline{\mu}} + A (x^*)^{\beta_1(\underline{\mu})} = \frac{D_{10} x^*}{r - \overline{\mu}} + A_1 (x^*)^{\beta_1(\overline{\mu})} + B (x^*)^{\beta_2(\overline{\mu})}.$$

This yields

$$A = \left(\frac{D_{10}}{r - \overline{\mu}} - \frac{D_{10}}{r - \underline{\mu}} + A_1 (x^*)^{\beta_1(\overline{\mu}) - 1} + B (x^*)^{\beta_2(\overline{\mu}) - 1} \right) (x^*)^{1 - \beta_1(\underline{\mu})}. \tag{4.19}$$

Now we can write

$$\tilde{L_1}(x) = \frac{D_{10}x}{r-\underline{\mu}} + \left(\left(\frac{D_{10}}{r-\overline{\mu}} - \frac{D_{10}}{r-\underline{\mu}} + \left[\frac{D_{11}}{r-\mu} - \frac{D_{10}}{r-\overline{\mu}} - B(x_2^F)^{\beta_2(\overline{\mu})-1}\right]\right.\right.$$
$$\left.\left.(x_2^F)^{1-\beta_1(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})-1} + B(x^*)^{\beta_2(\overline{\mu})-1}\right)(x^*)^{1-\beta_1(\underline{\mu})}\right)x^{\beta_1(\underline{\mu})}.$$

The smooth pasting condition at x^* yields that the first derivative of $\tilde{L_1}$ at x^* is equal to zero. Therefore,

$$\tilde{L_1}'(x) = \frac{D_{10}}{r-\underline{\mu}} + \beta_1(\underline{\mu})\left(\frac{D_{10}}{r-\overline{\mu}} - \frac{D_{10}}{r-\underline{\mu}} + \left[\frac{D_{11}}{r-\mu} - \frac{D_{10}}{r-\overline{\mu}} - B(x_2^F)^{\beta_2(\overline{\mu})-1}\right]\right.$$
$$\left.(x_2^F)^{1-\beta_1(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})-1} + B(x^*)^{\beta_2(\overline{\mu})-1}\right)$$
$$= 0.$$

This implies

$$-\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}} - \left(\frac{D_{10}}{r-\overline{\mu}} - \frac{D_{10}}{r-\underline{\mu}}\right) = \left[\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} - B(x_2^F)^{\beta_2(\overline{\mu})-1}\right]$$
$$\left.(x_2^F)^{1-\beta_1(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})-1} + B(x^*)^{\beta_2(\overline{\mu})-1}\right)$$
$$\iff \left(-\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}} - \left(\frac{D_{10}}{r-\overline{\mu}} - \frac{D_{10}}{r-\underline{\mu}}\right)\right)(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}$$
$$= \frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} - B(x_2^F)^{\beta_2(\overline{\mu})-1} + B(x^*)^{\beta_2(\overline{\mu})-1}(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}$$
$$\iff \left(-\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}} - \left(\frac{D_{10}}{r-\overline{\mu}} - \frac{D_{10}}{r-\underline{\mu}}\right)\right)(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}$$
$$-\frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}}$$
$$= B\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1} - (x_2^F)^{\beta_2(\overline{\mu})-1}\right).$$

Therefore, we get

$$B = \frac{1}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1} - (x_2^F)^{\beta_2(\overline{\mu})-1}} \Bigg[\Bigg(\Bigg(-\frac{1}{\beta_1(\underline{\mu})} + 1 \Bigg) \frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg) \tag{4.20}$$

$$(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})} - \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}} \Bigg].$$

Plugging this into equation (4.18) gives

$$A_1 = \Bigg(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} - \Bigg[\frac{1}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1} - (x_2^F)^{\beta_2(\overline{\mu})-1}} \tag{4.21}$$

$$\Bigg(\Bigg(-\frac{1}{\beta_1(\underline{\mu})} + 1 \Bigg) \frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg) (x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}$$

$$- \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}} \Bigg] (x_2^F)^{\beta_2(\overline{\mu})-1} \Bigg) (x_2^F)^{1-\beta_1(\overline{\mu})}.$$

Now, using (4.20) and (4.21), we get

$$\tilde{L}_1(x) =$$

$$\frac{D_{10}x}{r-\underline{\mu}} + \Bigg[\frac{D_{10}x^*}{r-\overline{\mu}} - \frac{D_{10}x^*}{r-\underline{\mu}} + \Bigg[\Bigg(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg) (x_2^F)^{1-\beta_1(\overline{\mu})}$$

$$- \frac{(x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1} - (x_2^F)^{\beta_2(\overline{\mu})-1}} \Bigg[\Bigg(\Bigg(1 - \frac{1}{\beta_1(\underline{\mu})} \Bigg) \frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg)$$

$$(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})} - \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}} \Bigg] \Bigg] (x^*)^{\beta_1(\overline{\mu})}$$

$$+ \frac{1}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{1-\beta_1(\overline{\mu})} - (x_2^F)^{\beta_2(\overline{\mu})-1}} \Bigg[\Bigg(\Bigg(-\frac{1}{\beta_1(\underline{\mu})} + 1 \Bigg) \frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg)$$

$$(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})} - \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}} \Bigg] (x^*)^{\beta_2(\overline{\mu})} \Bigg] \left(\frac{x}{x^*} \right)^{\beta_1(\underline{\mu})}$$

$$= \frac{D_{10}x}{r-\underline{\mu}} + \Bigg[\frac{D_{10}x^*}{r-\overline{\mu}} - \frac{D_{10}x^*}{r-\underline{\mu}} + \Bigg(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} \Bigg) (x_2^F)^{1-\beta_1(\overline{\mu})}(x^*)^{\beta_2(\overline{\mu})}$$

$$
-\frac{(x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\left[\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)\right.
$$

$$
\left.\left.(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}-\frac{D_{11}}{r-\underline{\mu}}+\frac{D_{10}}{r-\overline{\mu}}\right]\right]\left(\frac{x}{x^*}\right)^{\beta_1(\underline{\mu})}
$$

$$
=\frac{D_{10}x}{r-\underline{\mu}}+\left[\frac{\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}\right)}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}x^*\left(\frac{D_{10}}{r-\overline{\mu}}-\frac{D_{10}}{r-\underline{\mu}}\right)\right.
$$

$$
-\frac{\left((x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_2(\overline{\mu})}\right)(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}
$$

$$
\left.\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)\right]\left(\frac{x}{x^*}\right)^{\beta_1(\underline{\mu})}.
$$

Let's split the last expression into two parts. Firstly, we consider the part concerning $\frac{D_{10}}{r-\overline{\mu}}$:

$$
\begin{aligned}
&\frac{\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}\right)x^*}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{D_{10}}{r-\overline{\mu}}\\
&+\frac{\left((x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}x^*-(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})+1}\right)(x_2^F)^{\beta_1(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{D_{10}}{r-\overline{\mu}}\\
=&\frac{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})+1}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})+1}(x_2^F)^{\beta_1(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{D_{10}}{r-\overline{\mu}}\\
=&0.
\end{aligned}
$$

What remains is to consider the part concerning $\frac{D_{10}}{r-\underline{\mu}}$. That is:

$$
\begin{aligned}
&-\frac{\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}\right)x^*}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{D_{10}}{r-\underline{\mu}}\\
&-\frac{\left((x_2^F)^{\beta_2(\overline{\mu})-1}-(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}\right)x^*}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{D_{10}}{r-\underline{\mu}}\\
&+\frac{\left((x_2^F)^{\beta_2(\overline{\mu})-1}-(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}\right)x^*}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}}
\end{aligned}
$$

$$= -\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}}\frac{x^*\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}\right)}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}$$
$$= -\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}}x^*.$$

Therefore, we get

$$\tilde{L}_1(x) = \frac{D_{10}x}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{x}{x^*}\right)^{\beta_1(\underline{\mu})}.$$

Let's move on to the computation of $\hat{L}_1$. Using the expressions (4.21) and (4.20) we find:

$$\begin{aligned}
&\hat{L}_1(x) = \\
&\frac{D_{10}x}{r-\overline{\mu}} + \Bigg(\frac{D_{11}(x_2^F)^{1-\beta_1(\overline{\mu})}}{r-\underline{\mu}} - \frac{D_{10}(x_2^F)^{1-\beta_1(\overline{\mu})}}{r-\overline{\mu}} \\
&- \Bigg[\frac{1}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\Bigg(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\Bigg) \\
&\quad (x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})} - \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}}\Bigg](x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}\Bigg)x^{\beta_1(\overline{\mu})} \\
&+ \frac{1}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}\Bigg[\Bigg(\left(-\frac{1}{\beta_1(\underline{\mu})}+1\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\Bigg) \\
&\quad (x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})} - \frac{D_{11}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}}\Bigg]x^{\beta_2(\overline{\mu})}.
\end{aligned}$$

To make it more tractable, we again split this expression into two parts. Firstly, we put all expressions containing the term $(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}})$ together.

$$\begin{aligned}
&\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)(x_2^F)^{1-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})} \\
&+ \frac{\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)\left((x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})}+x^{\beta_2(\overline{\mu})}\right)}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}
\end{aligned}$$

$$=\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)(x_2^F)(x^*)^{\beta_1(\overline{\mu})}$$

$$\frac{\left((x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})}-(x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})}+(x_2^F)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})}+x^{\beta_2(\overline{\mu})}\right)}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x_2^F)^{\beta_2(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}}$$

$$=\frac{(x^*)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x_2^F.$$

In a similar way, we get for the remaining part that

$$\frac{\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}$$

$$-\frac{\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)(x_2^F)^{\beta_2(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}x_1^{\beta}(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})-\beta_1(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})-1}-(x_2^F)^{\beta_2(\overline{\mu})-1}}$$

$$=x^*\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)(x^*)^{\beta_1(\overline{\mu})-1}x_2^F$$

$$\frac{(x_2^F)^{\beta_1(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}-(x_2^F)^{\beta_2(\overline{\mu})-1}(x^*)^{1-\beta_1(\overline{\mu})}x^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}$$

$$=\frac{(x_2^F)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}-(x_2^F)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x^*.$$

Putting this together eventually yields the desired expression:

$$\hat{L_1}(x)=\frac{D_{10}x}{r-\overline{\mu}}+\frac{(x^*)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x_2^F$$

$$+\frac{(x_2^F)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})}-(x_2^F)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x^*.$$

□

4.7.2 Proof of Lemma (5)

In this section, we show that if the worst–case for the leader value is not always given by the worst possible trend, there exists a unique value x^* at which the worst–case changes from $\underline{\mu}$ to $\overline{\mu}$.

PROOF: The critical value x^* is found by applying the smooth pasting condition $\hat{L_1}(\overline{\mu}, x^*) = 0$. The first derivative of $\hat{L_1}$ is given by

$$\begin{aligned}\hat{L_1'}(\overline{\mu}, x) =& \frac{D_{10}}{r-\overline{\mu}} + \frac{\beta_1(\overline{\mu})(x^*)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})-1} - \beta_2(\overline{\mu})(x^*)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \\ & \left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x_2^F \\ & + \frac{\beta_2(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})-1} - \beta_1(\overline{\mu})(x_2^F)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \\ & \left(\left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x^*.\end{aligned}$$

In order to prove the existence of x^*, we will show that if $x^* \uparrow x_2^F$, $\hat{L_1'}(\overline{\mu}, x^*)$ becomes negative, and if $x^* \downarrow 0$, $\hat{L_1'}(\overline{\mu}, x^*)$ becomes positive.

We have

$$\begin{aligned}\hat{L_1'}(\overline{\mu}, x^*) =& \frac{D_{10}}{r-\overline{\mu}} + \frac{(\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))(x^*)^{\beta_1(\overline{\mu})+\beta_2(\overline{\mu})-1}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \\ & \left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right) x_2^F \\ & + \frac{\beta_2(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})}(x^*)^{\beta_2(\overline{\mu})} - \beta_1(\overline{\mu})(x_2^F)^{\beta_2(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}} \\ & \left(\left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right).\end{aligned}$$

Clearly, $\lim_{x^* \uparrow x_2^F} \hat{L_1'}(\overline{\mu}, x^*)$ has the same sign as the following expression.

$$\begin{aligned}& \frac{D_{10}}{r-\overline{\mu}}\left((x_2^F)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x_2^F)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}\right) \\ & + (\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))\,(x_2^F)^{\beta_1(\overline{\mu})+\beta_2(\overline{\mu})} \\ & \left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}} - \left(1 - \frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} + \frac{D_{10}}{r-\overline{\mu}}\right).\end{aligned} \tag{4.22}$$

Using the fact that $\frac{1}{\beta_1(\underline{\mu})} < \frac{D_{10}-D_{11}}{D_{10}}$ yields that expression (4.22) is smaller than

$$(\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))\,(x_2^F)^{\beta_1(\overline{\mu})+\beta_2(\overline{\mu})}\frac{1}{r-\underline{\mu}}\,(D_{11} - D_{10} + D_{10} - D_{11}) = 0. \quad (4.23)$$

Considering the case $x^* \downarrow 0$, one can easily see that $\lim_{x^*\downarrow 0} \hat{L_1'}(\overline{\mu}, x^*)$ has the same sign as

$$\begin{aligned}
&\lim_{x^*\downarrow 0}\Bigg(\frac{D_{10}}{r-\overline{\mu}}\left((x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}\right)\\
&\qquad + (\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))\,(x^*)^{\beta_1(\overline{\mu})+\beta_2(\overline{\mu})-1}\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F\\
&\qquad + \left(\beta_2(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})}(x^*)^{\beta_2(\overline{\mu})} - \beta_1(\overline{\mu})(x_2^F)^{\beta_2(\overline{\mu})}(x^*)^{\beta_1(\overline{\mu})}\right)\\
&\qquad \left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)\Bigg)\\
=&\lim_{x^*\downarrow 0}\Bigg((x^*)^{\beta_2(\overline{\mu})}\Bigg(\frac{D_{10}}{r-\overline{\mu}}\left((x_2^F)^{\beta_2(\overline{\mu})} - (x^*)^{\beta_1(\overline{\mu})-\beta_2(\overline{\mu})}\right)\\
&\qquad + (\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))(x^*)^{\beta_1(\overline{\mu})-1}\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F\\
&\qquad + \left(\beta_2(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})} - \beta_1(\overline{\mu})(x^*)^{\beta_1(\overline{\mu})-\beta_2(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}\right)\\
&\qquad \left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)\Bigg)\Bigg)\\
=&\lim_{x^*\downarrow 0}\underbrace{(x^*)^{\beta_2(\overline{\mu})}}_{\to+\infty}\Bigg(\frac{D_{10}}{\underbrace{r-\overline{\mu}}_{>0}}\left((x_2^F)^{\beta_2(\overline{\mu})} - \underbrace{(x^*)^{\beta_1(\overline{\mu})-\beta_2(\overline{\mu})}}_{\to 0}\right)\\
&\qquad + \underbrace{(\beta_1(\overline{\mu}) - \beta_2(\overline{\mu}))(x^*)^{\beta_1(\overline{\mu})-1}}_{\to 0}\left(\frac{D_{11}}{r-\underline{\mu}} - \frac{D_{10}}{r-\overline{\mu}}\right)x_2^F\\
&\qquad + \left(\underbrace{\beta_2(\overline{\mu})(x_2^F)^{\beta_1(\overline{\mu})}}_{<0} - \underbrace{\beta_1(\overline{\mu})(x^*)^{\beta_1(\overline{\mu})-\beta_2(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}_{\to 0}\right)
\end{aligned}$$

$$\underbrace{\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)}_{<0}\Bigg).$$

Therefore, we get $\hat{L_1'}(\overline{\mu}, x^*) > 0$ for x^* close to 0. Due to continuity of L_2' on $[0, x_2^F]$, we can find in that region a solution to $\hat{L_1'}(\overline{\mu}, x^*) = 0$.

The uniqueness of x^* is automatically given by the uniqueness of the solution to PDE (4.15).

□

4.7.3 Concavity of L_1

In this section, we prove that the leader function of the ambiguous firm is concave on $[0, x_2^F]$. In case the worst prior is always given by the lowest possible trend, this statement is trivial. The next proof shows that concavity is not lost even if the worst–case changes at some point.

PROOF: Suppose condition (4.11) is not satisfied (i.e. $\underline{\mu}$ is not always the worst–case). The concavity of $L_1(x)$ for $x < x^*$ is trivial. We therefore consider the second derivative of $L_1(x)$ in the interval $[x^*, x_2^F)$.

$$\hat{L_1''}(\overline{\mu}, x) =$$

$$\frac{\beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)(x^*)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})-2}-\beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)(x^*)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})-2}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}$$

$$\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x_2^F$$

$$+\frac{\beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)(x_2^F)^{\beta_1(\overline{\mu})}x^{\beta_2(\overline{\mu})-2}-\beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)(x_2^F)^{\beta_2(\overline{\mu})}x^{\beta_1(\overline{\mu})-2}}{(x^*)^{\beta_2(\overline{\mu})}(x_2^F)^{\beta_1(\overline{\mu})}-(x^*)^{\beta_1(\overline{\mu})}(x_2^F)^{\beta_2(\overline{\mu})}}$$

$$\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x^*.$$

Now, we have

$$\beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)x^{\beta_1(\overline{\mu})-2}\Bigg[\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x_2^F(x^*)^{\beta_2(\overline{\mu})}$$

$$
\begin{aligned}
&\qquad\qquad -\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x^*(x_2^F)^{\beta_2(\overline{\mu})}\Bigg]\\
&< \beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)x^{\beta_1(\overline{\mu})-2}x^*(x_2^F)^{\beta_2(\overline{\mu})}\\
&\quad\left[\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)-\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)\right]\\
&= \beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)x^{\beta_1(\overline{\mu})-2}x^*(x_2^F)^{\beta_2(\overline{\mu})}\left[\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\underline{\mu}}+\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}}\right]\\
&< \beta_1(\overline{\mu})(\beta_1(\overline{\mu})-1)x^{\beta_1(\overline{\mu})-2}x^*(x_2^F)^{\beta_2(\overline{\mu})}\frac{1}{r-\underline{\mu}}\left[D_{11}-D_{10}+D_{10}-D_{11}\right]\\
&= 0,
\end{aligned}
$$

where we used the fact that $x^*(x_2^F)^{\beta_2(\overline{\mu})} < (x^*)^{\beta_2(\overline{\mu})}(x_2^F)$ (because $x^* < x_2^F$ and $\beta_2(\overline{\mu}) < 0$) and $\frac{D_{10}-D_{11}}{D_{10}} > \frac{1}{\beta_1(\underline{\mu})}$.

In a similar way, we have

$$
\begin{aligned}
&\beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)x^{\beta_2(\overline{\mu})-2}\Bigg[-\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x_2^F(x^*)^{\beta_1(\overline{\mu})}\\
&\qquad\qquad +\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)x^*(x_2^F)^{\beta_1(\overline{\mu})}\Bigg]\\
&< \beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)x^{\beta_2(\overline{\mu})-2}x^*(x_2^F)^{\beta_1(\overline{\mu})}\\
&\left[-\left(\frac{D_{11}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)+\left(\left(1-\frac{1}{\beta_1(\underline{\mu})}\right)\frac{D_{10}}{r-\underline{\mu}}-\frac{D_{10}}{r-\overline{\mu}}\right)\right]\\
&= \beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)x^{\beta_2(\overline{\mu})-2}x^*(x_2^F)^{\beta_1(\overline{\mu})}\left[-\frac{D_{11}}{r-\underline{\mu}}+\frac{D_{10}}{r-\underline{\mu}}-\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}}{r-\underline{\mu}}\right]\\
&< \beta_2(\overline{\mu})(\beta_2(\overline{\mu})-1)x^{\beta_2(\overline{\mu})-2}x^*(x_2^F)^{\beta_1(\overline{\mu})}\frac{1}{r-\underline{\mu}}\left[-D_{11}+D_{10}-D_{10}+D_{11}\right]\\
&= 0,
\end{aligned}
$$

which proves the concavity of L_1.

□

4.7.4 Proof of Proposition (2)

The proof follows along very similar lines as the proof of Theorem (4). We take the same procedure, but now considering the value function in the continuation region, i.e. before any investment has taken place. Applying the BSDE approach and applying different value matching and smooth pasting conditions eventually yield the desired stopping time.

Proof: Denote

$$Y_t = \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \Big[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \Big| \mathcal{F}_t \Big].$$

Using the time consistency property of a rectangular set of density generators yields

$$\begin{aligned} Y_t &= \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \Big[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds \\ &\qquad + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \Big| \mathcal{F}_t \Big] \\ &= \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\inf_{Q' \in \mathscr{P}^\Theta} \mathsf{E}^{Q'} \left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + \int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-t)} D_{10} X_s ds \right.\right. \\ &\qquad \left.\left. + \int_{\tau_2^F}^{\infty} e^{-r(s-t)} D_{11} X_s ds \Big| \mathcal{F}_{\tau_{L,1}^t} \right] \Big| \mathcal{F}_t \right] \\ &= \inf_{Q \in \mathscr{P}^\Theta} \mathsf{E}^Q \left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds \right. \\ &\qquad + e^{-r(\tau_{L,1}^t - t)} \inf_{Q' \in \mathscr{P}^\Theta} \mathsf{E}^{Q'} \left[\int_{\tau_{L,1}^t}^{\tau_2^F} e^{-r(s-\tau_{L,1}^t)} D_{10} X_s ds \right. \\ &\qquad \left.\left. + \int_{\tau_2^F}^{\infty} e^{-r(s-\tau_{L,1}^t)} D_{11} X_s ds \Big| \mathcal{F}_{\tau_{L,1}^t} \right] \Big| \mathcal{F}_t \right] \end{aligned}$$

$$= \inf_{Q\in\mathscr{P}^\Theta} \mathsf{E}^Q\left[\int_t^{\tau_{L,1}^t} e^{-r(s-t)} D_{00} X_s ds + e^{-r(\tau_{L,1}^t - t)} L_1(x_{\tau_{L,1}^t}) \Big| \mathcal{F}_t\right].$$

Chen and Epstein (2002) show that Y_t solves the BSDE

$$-dY_t = g(Z_t)dt - Z_t dB_t,$$

for the *generator*

$$g(z) = -\kappa|z| - rY_t + X_t D_{00}.$$

The boundary condition is given by

$$Y_{\tau_{L,1}^t} = L(x_1^L),$$

where $L(x_1^L)$ is given by Theorem (4) and $x_1^L = x_{\tau_{L,1}^t}$.

Denote the present value of the leader payoff by Λ, i.e.

$$\Lambda(x_t) = Y_t.$$

The non–linear Feynman–Kac formula implies that Λ solves the non–linear PDE

$$\mathcal{L}_X \Lambda(x) + g(\sigma x \Lambda'(x)) = 0.$$

Hence, Λ solves

$$\frac{1}{2}\sigma^2 x^2 \Lambda''(x) + \mu x \Lambda'(x) - \kappa\sigma x \left|\Lambda'(x)\right| - r\Lambda(x) + D_{00} x = 0. \tag{4.24}$$

In the continuation region the leader function has to be increasing, hence $\Lambda' > 0$. This implies that $\underline{\mu}$ is the worst–case in the continuation region.

Therefore, equation (4.24) becomes

$$\begin{aligned}&\frac{1}{2}\sigma^2 x^2 \Lambda''(x) + (\mu - \kappa\sigma) x \Lambda'(x) - r\Lambda(x) + D_{00} x\\ =&\frac{1}{2}\sigma^2 x^2 \Lambda''(x) + \underline{\mu} x \Lambda'(x) - r\Lambda(x) + D_{00} x\\ =&0.\end{aligned}$$

The general increasing solution to this PDE is

$$\Lambda(x) = \frac{D_{00}x}{r-\underline{\mu}} + A_2 x^{\beta_1(\underline{\mu})}.$$

We have to distinguish two cases here. Either condition (4.11) holds, which means that the boundary condition takes the form (4.12) or the boundary condition becomes (4.13).

We will show that for both cases, the optimal investment threshold becomes

$$x_1^L = \frac{\beta_1(\underline{\mu})}{\beta_1(\underline{\mu})-1}\frac{I(r-\underline{\mu})}{D_{10}-D_{00}}. \tag{4.25}$$

If condition (4.11) is satisfied, the boundary condition is given by

$$L_1(x_1^L) = \frac{D_{10}x_1^L}{r-\underline{\mu}} + \left(\frac{x_1^L}{x_2^F}\right)^{\beta_1(\underline{\mu})} \frac{D_{11}-D_{10}}{r-\underline{\mu}} x_2^F - I.$$

Otherwise, the boundary condition is given by

$$L_1(x_1^L) = \frac{D_{10}x_1^L}{r-\underline{\mu}} - \frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{x_1^L}{x^*}\right)^{\beta_1(\underline{\mu})} - I.$$

In addition to the value matching condition, we apply a smooth pasting condition. Smooth pasting implies that the derivatives of the value function Λ and L coincide at $x_{\tau_{L,1}^t}$, i.e.

$$\Lambda'(x_{\tau_{L,1}^t}) = L_1'(x_{\tau_{L,1}^t}). \tag{4.26}$$

This condition ensures differentiability at the investment threshold.

Applying condition (4.26) gives

$$\frac{D_{00}}{r-\underline{\mu}} + \beta_1(\underline{\mu})A_2 {x_1^L}^{\beta_1(\underline{\mu})-1} = \frac{D_{10}}{r-\underline{\mu}} + \beta_1(\underline{\mu})A_1 {x_1^L}^{\beta_1(\underline{\mu})-1},$$

where

$$A_1 = \left(\frac{1}{x_2^F}\right)^{\beta_1(\underline{\mu})-1} \frac{D_{11}-D_{10}}{r-\underline{\mu}}$$

in the first case and

$$A_1 = -\frac{1}{\beta_1(\underline{\mu})}\frac{D_{10}x^*}{r-\underline{\mu}}\left(\frac{1}{x^*}\right)^{\beta_1(\underline{\mu})}$$

in the second.

Hence,

$$A_2 = \frac{D_{10}-D_{00}}{r-\underline{\mu}}\frac{1}{\beta_1(\underline{\mu})}\frac{1}{{x_1^L}^{\beta_1(\underline{\mu})-1}} + A_1.$$

Applying the value matching condition finally yields

$$\frac{D_{00}x_1^L}{r-\underline{\mu}} + \left(\frac{D_{10}-D_{00}}{r-\underline{\mu}}\frac{1}{\beta_1(\underline{\mu})}\frac{1}{{x_1^L}^{\beta_1(\underline{\mu})-1}} + A_1\right){x_1^L}^{\beta_1(\underline{\mu})} = \frac{D_{10}x_1^L}{r-\underline{\mu}} + A_1{x_1^L}^{\beta_1(\underline{\mu})} - I$$

$$\iff \frac{D_{10}-D_{00}}{r-\underline{\mu}}x_1^L - \frac{D_{10}-D_{00}}{r-\underline{\mu}}\frac{1}{\beta_1(\underline{\mu})}x_1^L = I$$

$$\iff \frac{\beta_1(\underline{\mu})-1}{\beta_1(\underline{\mu})}\frac{D_{10}-D_{00}}{r-\underline{\mu}}x_1^L = I,$$

and therefore, for both cases, it holds that

$$x_1^L = \frac{\beta_1(\underline{\mu})}{\beta_1(\underline{\mu})-1}\frac{I(r-\underline{\mu})}{D_{10}-D_{00}}.$$

□

4.7.5 Backward Stochastic Differential Equations

For the proof of our main theorem, it is crucial to express the leader value function by a backward stochastic differential equation (BSDE). For this reason we briefly summarize in this section some important results regarding BSDEs that have been obtained in the literature. We refer to El Karoui et al. (1997), and El Karoui and Mazliak (1997) for a detailed survey of this topic.

We start with considering the following stochastic differential equation (SDE)

$$dX_s = \sigma(X_s)dB_s + \mu(X_s)ds, \tag{4.27}$$

with initial condition $X_0 = x$. Here (B_s) denotes a $d-$dimensional Brownian motion and $\sigma : \mathbb{R}^n \mapsto \mathbb{R}^{n+d}$ and $\mu : \mathbb{R}^n \mapsto \mathbb{R}^n$ are given Lipschitz functions.

One can show that the integral form is given by

$$X_t = x + \int_0^t \sigma(X_s(\omega))dB_s(\omega) + \int_0^t \mu(X_s(\omega))ds.$$

This process satisfies the Ito formula. It holds for a smooth function f on $\mathbb{R}^n \times [0, \infty)$

$$df(X_t, t) = \partial_t f(X_t, t)dt + \nabla_x f(X_t, t)dX_t + \frac{1}{2}\sum_{i,j=1}^{n}(\sigma\sigma^*)_{ij} D_{x_i x_j} f(X_t, t)dt,$$

where σ^* denotes the transpose of σ. With the help of this formula, one can show that the solution to SDE (4.27) is a diffusion process with the infinitesimal generator

$$\mathcal{L} = \sum_{i=1}^{n} \mu_i(x) D_{x_i} + \frac{1}{2}\sum_{i,j=1}^{n}(\sigma(x)\sigma^*(x))_{ij} D_{x_i x_j}.$$

In many situations, however, one can find an inverse type of problem where a terminal condition at T is given (a good example is the leader value function (4.8) of the ambiguous firm). Starting with this terminal condition, one then wants to explore the process backward in time.

For this purpose, backward stochastic differential equations were introduced. Formally, a backward stochastic differential equation is given by

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \tag{4.28}$$

where Y and Z are unknown processes, the function g is called the generator of the BSDE and $\xi = Y_T \in L_P^2(\mathcal{F}_T)$ denotes the terminal condition.

In its differential form (4.28) can be written as

$$dY_s = -g(s, Y_s, Z_s)ds + Z_s dB_s, \quad s \in [0, T].$$

This type of differential equations were first studied for a linear generator by

Bismut (1973).

The following theorem gives an important existence and uniqueness result for the solution of BSDE (4.28).

Theorem 6 *(Pardoux and Peng (1990)) Let $M^2(0,T,\mathbb{R}^n)$ be the space of all $\mathbb{R}^n$-valued stochastic processes η on $[0,T]$ which satisfy $E\int_0^T |\eta_t|^2 < \infty$. Let g : $\Omega \times [0,\infty) \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ be a given function such that $g(\cdot, y, z) \in M^2(0,T,\mathbb{R}^m)$ for each T and for each fixed $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^{m\times d}$. Further, let g be a Lipschitz function, i.e. there exists a constant c such that*

$$|g(\omega,t,y,z) - g(\omega,t,y',z')| \leq c(|y-y'| + |z-z'|), \;\; y,y' \in \mathbb{R}^m, \;\; z,z' \in \mathbb{R}^{m\times d}.$$

Then, for each given $Y_T = \xi \in L^2_P(\mathcal{F}_t)$, there exists a unique pair of processes $(Y,Z) \in M^2(0,T,\mathbb{R}^m \times \mathbb{R}^{m\times d})$ satisfying BSDE (4.28). Moreover, Y has continuous path, a.s.

Hence, for suitable conditions stipulated on the generator g, there exists always a unique solution to BSDE (4.28).

For the one dimensional case $m = 1$, one can obtain the following comparison theorem.

Theorem 7 *(Peng (1992)) Let the same condition for two generators g_1 and g_2 hold as in the last theorem. Further let $m = 1$. If $\xi_1 \geq \xi_2$ a.s. and $g_1 \geq g_2$ a.s., then for the correspondent processes it holds that*

$$Y_t^1 \geq Y_t^2 \;\; a.s.$$

for any time t.

That means, for comparison in the one dimensional case it suffices to consider the generator and the terminal conditions of the BSDEs.

Finally, an important result concerns the relationship between backward stochastic differential equations and partial differential equations (PDEs). Often times it is much more convenient to solve a PDE rather than a BSDE (like it was the case in the proof of Theorem (4)). The next theorem describes the connection between a solution of a particular PDE and a BSDE.

Suppose $X_s^{t,x}, s \in [t,T]$ is a solution to SDE (4.27) with initial condition $X_t^{t,x} = x \in \mathbb{R}^n$. Consider the following BSDE

$$dY_s^{t,x} = -g(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})ds + Z_s^{t,x} dB_s, \tag{4.29}$$

with $Y_T^{t,x} = \phi(X_T^{t,x})$. Consider the case $m = 1$. Then the following theorem was first established by Peng (1991).

Theorem 8 *Assume that b, σ, and ϕ are given Lipschitz functions on $\mathbb{R}^n$ that take values in $\mathbb{R}^n$, $\mathbb{R}^{n\times d}$ and $\mathbb{R}$, respectively, and that g is a real valued Lipschitz function on $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d$. Then the following relation holds*

$$Y_s^{t,x} = u(s, X_s^{t,x}).$$

In particular, we have $u(t,x) = Y_t^{t,x}$, where $u = u(t,x)$ is the unique viscosity solution of the following PDE

$$\partial_t u + \mathcal{L}u + g(x, u, \sigma^* Du) = 0,$$

where $Du = (D_{x_1}u, ..., D_{x_n}u)$ and with terminal condition $u|_{t=T} = \phi$.

Bibliography

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., and Ku, H. (2007). Coherent multiperiod risk adjusted values and bellmans principle. *Annals of Operations Research*, 152(1):5–22.

Aumann, R. J. and Serrano, R. (2008). An economic index of riskiness. *Journal of Political Economy*, 116(5):810–836.

Azevedo, A. and Paxson, D. (2014). Developing real option game models. *European Journal of Operational Research*, 237(3):909–920.

Bali, T. G., Cakici, N., and Chabi-Yo, F. (2011). A generalized measure of riskiness. *Management Science*, 57(8):1406–1423.

Bewley, T. F. (2002). Knightian decision theory part i. *Decisions in Economics and Finance*, 25(2):79–110.

Bismut, J.-M. (1973). Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Boyarchenko, S. and Levendorski, S. (2014). Preemption games under levy uncertainty. *Games and Economic Behavior*, 88(0):354–380.

Boyer, M., Lasserre, P., Mariotti, T., and Moreaux, M. (2004). Preemption and rent dissipation under price competition. *International Journal of Industrial Organization*, 22(3):309–328.

Chen, Z. and Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443.

Cheng, X. and Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics*, 7(1):29–68.

Cheridito, P., Delbaen, F., and Kupper, M. (2004). Dynamic monetary risk measures for bounded discrete-time processes. *Working Paper, ETH Zrich.*

Cheridito, P. and Kupper, M. (2011). Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(01):137–162.

Cheridito, P. and Stadje, M. (2009). Time-inconsistency of var and time-consistent alternatives. *Finance Research Letters*, 6(1):40–46.

de Villemeur, E. B., Ruble, R., and Versaevel, B. (2014). Investment timing and vertical relationships. *International Journal of Industrial Organization*, 33(0):110–123.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. *Advances in finance and stochastics*, pages 1–37.

Detlefsen, K. and Scandolo, G. (2005). Conditional and dynamic convex risk measures. *Finance and Stochastics*, 9(4):539–561.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under uncertainty.* Princeton Univ. Press, Princeton, NJ.

Einhorn, D. (2008). Private profits and socialized risk. *Global Association of Risk Professionals*, 42:10–18.

El Karoui, N. and Mazliak, L. (1997). *Backward stochastic differential equations.* Longman, Harlow.

El Karoui, N., Peng, S., and Quenez, M. (1997). Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75:643–669.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447.

Föllmer, H. and Schied, A. (2004). *Stochastic finance. An introduction in discrete time*. De Gruyter studies in mathematics ; 27. de Gruyter, Berlin [u.a.].

Föllmer, H. and Schied, A. (2011). *Stochastic finance. An introduction in discrete time*. de Gruyter, Berlin [u.a.].

Foster, D. and Hart, S. (2009). An operational measure of riskiness. *Journal of Political Economy*, 117:785–814.

Foster, D. P. and Hart, S. (2013). A wealth-requirement axiomatization of riskiness. *Theoretical Economics*, 8(2):591–620.

Frittelli, M. and Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486.

Fudenberg, D. and Tirole, J. (1985). Preemption and rent equilization in the adoption of new technology. *Review of Economic Studies*, 52(3):383–401.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Hart, S. (2011). Comparing risks by acceptance and rejection. *Journal of Political Economy*, 119(4):617–638.

Hellmann, T. and Riedel, F. (2015). A dynamic extension of the foster-hart measure of riskiness. *Journal of Mathematical Economics*, 59:66–70.

Hellmann, T. and Thijssen, J. J. (2015). Ambiguity in a real option game. *Working Paper 533, Center for Mathematical Economics, Bielefeld University.*

Homm, U. and Pigorsch, C. (2012). An operational interpretation and existence of the aumann-serrano index of riskiness. *Economics Letters*, 114(3):265–267.

Hoppe, H. C. (2000). Second-mover advantages in the strategic adoption of new technology under uncertainty. *International Journal of Industrial Organization*, 18(2):315–338.

Huisman, K. (2001). *Technology Investment: A Game Theoretic Real Options Approach*. Klumer Academic Publishers, Dordrecht, Netherlands.

Huisman, K. and Kort, P. (1999). Effects of strategic interactions on the option value of waiting. *Working Paper CentER DP no. 9992, Tilburg University, Tilburg.*

Huisman, K. J. and Kort, P. M. (2015). Strategic capacity investment under uncertainty. *The RAND Journal of Economics*, 46(2):376–408.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions*. Wiley, New York [u.a.].

Karatzas, I. and Cvitanic, J. (1999). On dynamic measures of risk. *Finance and Stochastics*, 3(4):451–482.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Knight, F. H. (1921). *Risk, Uncertainty and Profit.* Houghton Mifflin Co, Boston, MA.

Kort, P. and Trojanowska, M. (2010). The worst case for real options. *Journal of Optimization Theory and Applications*, 146:709–734.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.

Mason, R. and Weeds, H. (2010). Investment, uncertainty and pre-emption. *International Journal of Industrial Organization*, 28(3):278–287.

McDonald, R. and Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics*, 101(4):707–27.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):pp. 141–183.

Murto, P. (2004). Exit in duopoly under uncertainty. *RAND Journal of Economics*, 35(1):111–127.

Nishimura, K. G. and Ozaki, H. (2007). Irreversible investment and knightian uncertainty. *Journal of Economic Theory*, 136(1):668–694.

Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61.

Pawlina, G. and Kort, P. M. (2006). Real options in an asymmetric duopoly: Who benefits from your competitive disadvantage? *Journal of Economics & Management Strategy*, 15(1):1–35.

Peng, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports*, 37(1-2):61–74.

Peng, S. (1992). A generalized dynamic programming principle and hamilton-jacobi-bellman equation. *Stochastics and Stochastic Reports*, 38(2):119–134.

Peng, S. (1997). Backward sde and related g-expectation. *Backward Stochastic Differential Equation, Pitman res. Notes Math. Ser., 364, Longman, Harlow*, 141–159.

Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200.

Riedel, F. (2013). *Die Schuld der Ökonomen : was Ökonomie und Mathematik zur Krise beitrugen*. Econ, Berlin.

Riedel, F. and Hellmann, T. (2015). The foster-hart measure of riskiness for general gambles. *Theoretical Economics*, 10(1):1–9.

Riedel, F. and Steg, J.-H. (2014). Subgame-perfect equilibria in stochastic timing games. *Working Paper 524, Center for Mathematical Economics, Bielefeld University*, (524).

Roorda, B., Schumacher, J. M., and Engwerda, J. (2005). Coherent Acceptability Measures In Multiperiod Models. *Mathematical Finance*, 15(4):589–612.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.

Schreiber, A. (2012). An economic index of relative riskiness. Discussion paper series, The Center for the Study of Rationality, Hebrew University, Jerusalem.

Shiryaev, A. N. (1984). *Probability*. Graduate texts in mathematics ; 95. Springer, New York.

Smets, F. (1991). Exporting versus fdi: The effect of uncertainty, irreversibilities and strategic interactions. *Working Paper, Yale University, New Haven.*

Steg, J.-H. and Thijssen, J. J. (2015). Quick or persistent? strategic investment demanding versatility. *Working Paper 541, Center for Mathematical Economics, Bielefeld University.*

Thijssen, J. J. (2010). Preemption in a real option game with a first mover advantage and player-specific uncertainty. *Journal of Economic Theory*, 145(6):2448–2462.

Thijssen, J. J. (2011). Incomplete markets, ambiguity, and irreversible investment. *Journal of Economic Dynamics and Control*, 35(6):909–921.

Thijssen, J. J., Huisman, K., and Kort, P. (2012). Symmetric equilibrium strategies in game theoretic real option models. *Journal of Mathematical Economics*, 48(4):219–225.

Wang, T. (1999). A class of dynamic risk measures. *Working Paper, University of British Columbia.*

Weber, S. (2006). Distribution-Invariant Risk Measures, Information, And Dynamic Consistency. *Mathematical Finance*, 16(2):419–441.

Weeds, H. (2002). Strategic Delay in a Real Options Model of R&D Competition. *Review of Economic Studies*, 69(3):729–47.